Query Match 4.0%; Score 91; DB 14; Length 522;
Best Local Similarity 63.5%; Pred. No. 6e-12;
Matches 139; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy       1397 AGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAGAAAGCG 1456
              |||  |||||  | |  ||||| |||||||  | ||  |  ||  | || ||||||   |
Db         23 AGGAAGAAGAAGAGGAAGAGGAAGAAGAGGAGGAAGAAGAAGAGGAAGAGGAAGAAGAGG 82

Qy       1457 AAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGGATAATG 1516
              | || ||||  |  ||||| ||||  |  || ||||| |||       || || ||  | |
Db         83 AGGAAGAAGAAGAGGAAGAGGAAGAAGAGGAGGAAGAAGAAGAGGAAGAGGAAGAAGAGG 142

Qy       1517 GCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCGAAGAGG 1576
              |||||| |||||  | |   |  |||||| ||    |||||||  || ||  |  |||||||
Db        143 AGGAAGAAGAAGAGGAAGAGGAAGAAGAGGAGGAAGAAGAAGAGGAAGAGGAAGAAGAGG 202

Qy       1577 AGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
              ||||||  |||||    || ||| ||  |  || |||||
Db        203 AGGAAGAAGAAGAGGAAGAGGAAGAAGAGGAGGAAGAAG 241
```

Search completed: August 25, 2004, 14:58:18
Job time : 1146.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 11:49:00 ; Search time 5161.56 Seconds
(without alignments)
13086.780 Million cell updates/sec

Title: US-10-735-098-7
Perfect score: 2262
Sequence: 1 atgtgtaaaccgaattatgg.....tgcaggaggtggaaaaatga 2262

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		103.8	4.6	908	28	AZ548467	AZ548467 ENTEK30TR
2	c	101.8	4.5	434	14	CA745444	CA745444 wri2s.pk0
3	c	100.8	4.5	263	13	BU005139	BU005139 QGG7D20.y
4		100.4	4.4	395	29	CE760553	CE760553 tigr-gss-
5	c	100.4	4.4	843	28	AZ551618	AZ551618 ENTDV54TR
6	c	100.2	4.4	232	29	CG321416	CG321416 OGYBI53TV
7	c	100.2	4.4	233	12	BJ454017	BJ454017 BJ454017
8	c	98.8	4.4	265	14	CA607797	CA607797 wr1.pk008
9	c	98.8	4.4	880	28	AZ529191	AZ529191 ENTBV68TR
10		98.6	4.4	264	13	BQ241566	BQ241566 TaE05003E
11		98.6	4.4	332	14	CA566759	CA566759 K0405E07-
12		98.6	4.4	575	10	BF707975	BF707975 A634_LE A
13		98.4	4.4	960	28	CC143364	CC143364 NDL.32A13
14		98.2	4.3	594	12	BM167059	BM167059 EST569582
15		98.2	4.3	612	28	BZ149404	BZ149404 CH230-395
16	c	98.2	4.3	739	28	BZ220379	BZ220379 CH230-312
17		98.2	4.3	799	28	BZ201507	BZ201507 CH230-445
18	c	98.2	4.3	1096	14	CK206118	CK206118 FGAS01769
19	c	97.6	4.3	605	28	BZ222147	BZ222147 CH230-361
20		97.6	4.3	665	13	BU432980	BU432980 603221001
21		97.4	4.3	341	29	CE510654	CE510654 tigr-gss-
22		97	4.3	278	13	BQ460618	BQ460618 HA06H01r
23		97	4.3	617	14	CA743799	CA743799 wri1s.pk0
24	c	97	4.3	691	28	BH040136	BH040136 RPCI-24-2
25		97	4.3	890	28	BH146886	BH146886 ENTPK48TF
26	c	97	4.3	906	28	BH153606	BH153606 ENTTS83TF
27		96.6	4.3	531	14	CD915362	CD915362 G550.125O
28	c	96.6	4.3	546	28	AZ600777	AZ600777 1M0418D18
29	c	96.2	4.3	289	14	CA716220	CA716220 wdk3c.pk0
30		96.2	4.3	340	13	BQ462203	BQ462203 HD02A20r
31		96.2	4.3	471	12	BM166650	BM166650 EST569173
32		96.2	4.3	477	14	CD311846	CD311846 StrPu691.
33		96	4.2	622	13	BU418103	BU418103 603959413
34		95.8	4.2	302	9	AA572217	AA572217 vl51e08.r
35	c	95.4	4.2	311	14	CA699402	CA699402 wlk8.pk00
36	c	95.4	4.2	406	28	AZ509483	AZ509483 1M0352O19
37		95.4	4.2	924	13	BQ925195	BQ925195 AGENCOURT
38	c	95.2	4.2	548	9	AI746770	AI746770 ul07c01.y
39		95.2	4.2	675	28	BZ104011	BZ104011 CH230-227
40	c	95.2	4.2	3743	11	AK048184	AK048184 Mus muscu
41		94.4	4.2	537	14	CF380795	CF380795 lac78f12.
42		94.2	4.2	223	13	CA022136	CA022136 HZ42E07r
43		94.2	4.2	866	28	AZ527885	AZ527885 ENTCA79TR
44	c	94.2	4.2	877	28	AZ531291	AZ531291 ENTBQ34TR
45		94.2	4.2	912	28	AZ551092	AZ551092 ENTFJ22TF

ALIGNMENTS

RESULT 1
AZ548467
LOCUS AZ548467 908 bp DNA linear GSS 14-NOV-2000
DEFINITION ENTEK30TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION AZ548467
VERSION AZ548467.1 GI:11172102
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 908)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica HM1:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library
Seq primer: M13-Reverse
Class: shotgun

High quality sequence start: 17
High quality sequence stop: 828.

FEATURES Location/Qualifiers
source 1. .908
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site_1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell, Oxford University Press, 1999)."

ORIGIN

Query Match 4.6%; Score 103.8; DB 28; Length 908;
Best Local Similarity 45.5%; Pred. No. 2.9e-10;
Matches 369; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

```
Qy       1397 AGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAGAAAGCG 1456
Db         50 AAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGATGATGAAGAAGATGATG 109

Qy       1457 AAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGGATAATG 1516
Db        110 AAGAAGATGATGAAGAAGAAGAAGATGATGAAGATGAAGATGATGAAGACGATGAAGACG 169

Qy       1517 GCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCGAAGAGG 1576
Db        170 ACGAATATGAATTAGAAGATGATGATGATGAAGAAGAAGAAGATGATGATGAAGATGATG 229

Qy       1577 AGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAGGCAACGGCGGTTCAGGCAGCA 1636
Db        230 AAGACGATGAAGACGACGAATATGAATTAGAAGATGATGATGATGATGAAGAAGAAGATG 289

Qy       1637 TCCTGCCTGCCCTAGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCC 1696
Db        290 ATGATGAAGAAGAAGAAGATGATGATGAAGATGATGATGAAGAAGAAGATGAAGAAGATG 349

Qy       1697 GCACGGCAGAAACGGATATTCCGCAAAGCGGAACGGCGCATTATACCGGCACTTGGGAAG 1756
Db        350 ATGAAGAAGATGATGAAGATGATGATGAAGAAGAAGAAGATGATGATGAAGATGATGAAG 409

Qy       1757 CGCGTATCGGCAAACCCATTCAATGGGACAATCAGGCGGATGAAAAAGCGGCAAAAGCAG 1816
Db        410 ACGATGAAGACGACGAATATGAATTAGAAGATGATGATGAAGAAGAAGAAGATGATGATG 469

Qy       1817 AATTTACCGTTGATTTCGACAAGAAATCGATTTCCGGAAAGCTGACGGAGCAAAACGGCG 1876
Db        470 AAGAAGAAGAAGATGATGATGAAGATGATGATGAAGAAGAAGATGAAGATGATGATGAAG 529

Qy       1877 TAGAACCTGCTTTCCATATTGAAGACGGCAAGATTGATGGCAACGGTTTCCACGCGACAG 1936
Db        530 AAGAAGATGAAGATGATGAAGAAGAAGATGATGATGAAGAAGAAGATGAAGAAGAAG 589

Qy       1937 CGCGCACTCGGGAGAGCGGCATCAATCTTTCGGGAAATGGTTCGACCGACCCCAAAACAT 1996
Db        590 AAGATGATGAAGAAGATGATGAAGAAGAAGAAGATGATGATGAAGACGACGAATATGAAT 649

Qy       1997 TCCAAGCTAGTAATCTTCGTGTAGAAGGAGGATTTTACGGCCCGCAGGCGGCGGAATTGG 2056
Db        650 TAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGAAGATGATGAAGAAGAAG 709

Qy       2057 GCGGTACTATTTTCAATAATGATGGGAAATCTCTTAGTATAACTGAAAATATTGAAAATG 2116
```

```
Db        710 AAGATGATGATGATGATGATGAAGAAGATGATGATGAAGATGATGAAGAAGATGAAGATG 769

Qy       2117 AAGCTGAAGCTGAAGTTGAAGTTGAAGCTGAAGCTGAAGTTGAAGTTGAAGCTGATGTTG 2176
Db        770 ATGAAGATGATGAAGAAGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAG 829

Qy       2177 GCAAACAGTTAGAACCTGATGAAGTTAAACA 2207
Db        830 ATGATGAAGACGACGAATTTGAATTTAAAGA 860
```

RESULT 2
CA745444/c
LOCUS CA745444 434 bp mRNA linear EST 26-NOV-2002
DEFINITION wri2s.pk001.i6 wri2s Triticum aestivum cDNA clone wri2s.pk001.i6 5' end, mRNA sequence.
ACCESSION CA745444
VERSION CA745444.1 GI:25561267
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 434)
AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z., Miao,G., Caraher,N. and Hanafey,M.K.
TITLE DuPont Wheat cDNA Sequence
JOURNAL Unpublished (2002)
COMMENT Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: T7.
FEATURES Location/Qualifiers
source 1. .434
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wri2s.pk001.i6"
/tissue_type="leaf"
/lab_host="DH10B"
/clone_lib="wri2s"
/note="Vector: PGEM-T Easy; Site_1: SmaI; Riband (susceptible) wheat leaves infected with Septoria tritici strain A, 48 hours after infection, subtracted w/ comparable uninfected leaves"

ORIGIN

Query Match 4.5%; Score 101.8; DB 14; Length 434;
Best Local Similarity 66.2%; Pred. No. 6.5e-10;
Matches 145; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```
Qy       1397 AGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAGAAAGCG 1456
Db        375 AAGAGGAAGAGGAAGAGGAGGAAGAAGAAGAAGAAGNNGAGGAAGAAGAGGAGGAAGAAG 316

Qy       1457 AAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGGATAATG 1516
Db        315 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 256

Qy       1517 GCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCGAAGAGG 1576
Db        255 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 196

Qy       1577 AGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
Db        195 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 157
```

RESULT 3
BU005139/c
LOCUS BU005139 263 bp mRNA linear EST 22-AUG-2002
DEFINITION QGG7D20.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone QGG7D20, mRNA sequence.
ACCESSION BU005139
VERSION BU005139.1 GI:22439534
KEYWORDS EST.
SOURCE Lactuca sativa
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae; Lactuca.
REFERENCE 1 (bases 1 to 263)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenomics.ucdavis.edu/
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
singleton, see http://cgpdb.ucdavis.edu/ for details.
Plate: QGG7 row: D column: 20.

FEATURES Location/Qualifiers
source 1. .263
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="QGG7D20"
/lab_host="E.coli"
/clone_lib="QG_EFGHJ lettuce serriola"
/note="Vector: pBRcDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/ TAG_SEQ=Not found"

ORIGIN

Query Match 4.5%; Score 100.8; DB 13; Length 263;
Best Local Similarity 66.7%; Pred. No. 9.4e-10;
Matches 144; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy       1391 AACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAG 1450
             ||  | | |  |||||  |||  || || ||||| |  | ||  |  ||  | || ||||
Db        227 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 168

Qy       1451 AAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGG 1510
             ||    ||||| ||||  | ||||||  ||||  |  ||||||||  |||      |||||
Db        167 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 108

Qy       1511 ATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCG 1570
             |  | |  ||||| |||||||  |    | ||||||||||  ||||||||  || ||||  |
Db        107 AAGAAGAAGAAGAAGAAGAAGAAGAAGACGAAGAAGAAGACGAAGAAGAAGAAGAAGAAG 48

Qy       1571 AAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGC 1606
             ||||  || ||||  ||||||   ||||||  |  |||
Db         47 AAGAAGAAGAAGAAGAAGAAGAAGAAGACCCGC 12
```

RESULT 4
CE760553
LOCUS CE760553 395 bp DNA linear GSS 30-SEP-2003
DEFINITION tigr-gss-dog-17000369997532 Dog Library Canis familiaris genomic, genomic survey sequence.
ACCESSION CE760553
VERSION CE760553.1 GI:37101172
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 395)
AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES Location/Qualifiers
source 1. .395
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site_1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 4.4%; Score 100.4; DB 29; Length 395;
Best Local Similarity 63.0%; Pred. No. 1.2e-09;
Matches 155; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```
Qy       1391 AACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAG 1450
             ||  | | |  |||||  |||  || || ||||| || | ||| |  ||  | || ||||
Db         83 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGGAGAAGGAGAAGAAGAAGAAG 142

Qy       1451 AAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGG 1510
             ||    ||||| ||||  | ||||||  ||||  |  ||||||||  |||      |||||
Db        143 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGGAGAAGAAGAAG 202

Qy       1511 ATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCG 1570
             |  | |  ||||| |||||||  |    | ||||||| |   | |||||  || ||||  |
Db        203 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGGAGAAGGAGAAGAAGAAGAAGAAG 262

Qy       1571 AAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAGGCAACGGCGGTTCAG 1630
             |||| || ||||  ||||||   ||||||  ||       | || |||    || ||   ||
Db        263 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGGAGGAGGAGGAGGAGGAAGAG 322

Qy       1631 GCAGCA 1636
             |  | |
Db        323 GAGGAA 328
```

RESULT 5
AZ551618/c
LOCUS AZ551618 843 bp DNA linear GSS 14-NOV-2000
DEFINITION ENTDV54TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION AZ551618
VERSION AZ551618.1 GI:11176919
KEYWORDS GSS.

```
SOURCE      Entamoeba histolytica
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 843)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 39
            High quality sequence stop: 838.
FEATURES             Location/Qualifiers
     source          1. .843
                     /organism="Entamoeba histolytica"
                     /mol_type="genomic DNA"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
ORIGIN

  Query Match             4.4%;  Score 100.4;  DB 28;  Length 843;
  Best Local Similarity   45.4%;  Pred. No. 1.4e-09;
  Matches  359;  Conservative    0;  Mismatches  431;  Indels    0;  Gaps    0;

Qy       1396 AAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAGAAAGC 1455
             ||| |   ||| | |  || ||||| || |  | ||  | |||| | || |||||
Db        829 AAGAAGATTGATGATGAAGAAGATGATGATGAAGAAGATGATGATGATGAAGAAGATGAT 770

Qy       1456 GAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGGATAAT 1515
             ||||| || |  | |||||| ||    |  ||||| || ||    |   || || ||  |
Db        769 GAAGAAGATGATGAAGAAGAAGAGTATGATGAAGATGAAGATGATGAAGACGATGAAGAC 710

Qy       1516 GGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCGAAGAG 1575
             | |||| | |||  || |   | || || |||   ||||||| |||||| |  || ||
Db        709 GACGAATATGAATTAGAAGATGATGATGATGAAGAAGAAGAAGATGATGATGAAGATGAT 650

Qy       1576 GAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAGGCAACGGCGGTTCAGGCAGC 1635
             || || | ||||||   |||| |  |||     |  || |   | |  |    ||
Db        649 GAAGACGATGAAGACGACGAATATGAATTAGAAGATGATGATGATGATGAAGAAGAAGAT 590

Qy       1636 ATCCTGCCTGCCCTAGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATC 1695
                  |  |||||  | | |     ||    ||              |||
Db        589 GATGATGAAGAAGAAGAAGATGATGATGAAGATGATGATGAAGAAGAAGATGAAGAAGAT 530

Qy       1696 CGCACGGCAGAAACGGATATTCCGCAAAGCGGAACGGCGCATTATACCGGCACTTGGGAA 1755
              | |||    ||  |   |  | |   |  || ||   |   |     |||
Db        529 GATGAAGAAGATGATGAAGATGATGATGAAGAAGAAGAAGATGATGATGAAGATGATGAA 470

Qy       1756 GCGCGTATCGGCAAACCCATTCAATGGGACAATCAGGCGGATGAAAAAGCGGCAAAAGCA 1815
              |   |  | | |     | |||  ||  || | |  || ||| |||  |   | |
```

```
Db        469 GACGATGAAGACGACGAATATGAATTAGAAGATGATGATGAAGAAGAAGAAGATGATGAT 410

Qy       1816 GAATTTACCGTTGATTTCGACAAGAAATCGATTTCCGGAAAGCTGACGGAGCAAAACGGC 1875
              |||      | |||   ||  |  |      |  | | |     ||  |  | |
Db        409 GAAGAAGAAGAAGATGATGATGAAGATGATGATGAAGAAGAAGATGAAGATGATGATGAA 350

Qy       1876 GTAGAACCTGCTTTCCATATTGAAGACGGCAAGATTGATGGCAACGGTTTCCACGCGACA 1935
              | |||| ||       ||  |||||| |    |  |||||   | |     | |    |
Db        349 GAAGAAGATGAAGATGATGATGAAGAAGAAGATGATGATGAAGAAGAAGATGAAGAAGAA 290

Qy       1936 GCGCGCACTCGGGAGAGCGGCATCAATCTTTCGGGAAATGGTTCGACCGACCCCAAAACA 1995
              |      | ||    |      |     |  ||| |     ||||    |   |
Db        289 GAAGATGATGAAGAAGATGATGAAGAAGAAGAAGATGATGATGAAGACGACGAATATGAA 230

Qy       1996 TTCCAAGCTAGTAATCTTCGTGTAGAAGGAGGATTTTACGGCCCGCAGGCGGCGGAATTG 2055
              ||  ||| |  | |      ||  |||| ||    | | |       | |  |  |||
Db        229 TTAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGAAGATGATGAAGAAGAA 170

Qy       2056 GGCGGTACTATTTTCAATAATGATGGGAAATCTCTTAGTATAACTGAAAATATTGAAAAT 2115
              | | |  |  |    || |||| |  |  |  |                ||||  |  |||| ||
Db        169 GAAGATGATGATGATGATGATGAAGAAGATGATGATGAAGATGATGAAGAAGATGAAGAT 110

Qy       2116 GAAGCTGAAGCTGAAGTTGAAGTTGAAGCTGAAGCTGAAGTTGAAGTTGAAGCTGATGTT 2175
              || |  |||| ||||| ||| |  || | ||||| ||| |  || | |||| |||||
Db        109 GATGAAGAAGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAA 50

Qy       2176 GGCAAACAGT 2185
              | | |  | |
Db         49 GACGACGAAT 40


RESULT 6
CG321416/c
LOCUS       CG321416                232 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OGYBI53TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0640J09,
            genomic survey sequence.
ACCESSION   CG321416
VERSION     CG321416.1  GI:34238682
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 232)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGYBI53TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .232
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0640J09"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             4.4%;  Score 100.2;  DB 29;  Length 232;
  Best Local Similarity   65.3%;  Pred. No. 1.2e-09;
```

```
  Matches  147;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy       1391 AACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAG 1450
             ||  | | |  |||||| |||  || || ||||| |  | ||  |  ||  | || ||||
Db        226 AAGAAGAAGAAGAAGATGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 167

Qy       1451 AAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGG 1510
             ||   ||||| ||||  | |||||| |||| |  ||||||| |||      ||||| |
Db        166 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 107

Qy       1511 ATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCG 1570
             | | |  ||||| |||||||  |   |  |||||||||   |||||||  || ||||  |
Db        106 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 47

Qy       1571 AAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
             |||| || ||||  ||||||   ||||||  ||      | |||||
Db         46 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 2


RESULT 7
BJ454017/c
LOCUS       BJ454017                233 bp    mRNA    linear   EST 23-MAY-2002
DEFINITION  BJ454017 K. Sato unpublished cDNA library, cv. Akashinriki
            vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
            baak45j15 5', mRNA sequence.
ACCESSION   BJ454017
VERSION     BJ454017.1  GI:21132592
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 233)
  AUTHORS   Sato,K., Saisho,D. and Takeda,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .233
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="Akashinriki"
                     /sub_species="vulgare"
                     /db_xref="taxon:112509"
                     /clone="baak45j15"
                     /tissue_type="leaves"
                     /dev_stage="vegetative stage"
                     /clone_lib="K. Sato unpublished cDNA library, cv.
                     Akashinriki vegetative stage leaves"
ORIGIN

  Query Match             4.4%;  Score 100.2;  DB 12;  Length 233;
  Best Local Similarity   65.3%;  Pred. No. 1.2e-09;
  Matches  147;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy       1391 AACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAG 1450
             ||  | | |  |||||  |||  || || ||||| |  | ||  |  ||  | || ||||
Db        228 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 169

Qy       1451 AAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGG 1510
             ||   ||||| ||||  | |||||| || |  | |||||||||  ||||      ||||| |
Db        168 AAGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGACGAAGAAGAAGAAAAAGAAGAAGAAG 109

Qy       1511 ATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCG 1570
             | | |  ||||| |||||||  |   |  |||||||||   |||||||  || ||||  |
Db        108 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 49

Qy       1571 AAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
             |||| || ||||  ||||||   ||||||  ||      | |||||
Db         48 AAGATGAAGATGAAGAAGAAGAAGAAGAACAAGAACAAGAAGAAG 4


RESULT 8
CA607797/c
LOCUS       CA607797                265 bp    mRNA    linear   EST 21-NOV-2002
DEFINITION  wr1.pk0082.f4 wr1 Triticum aestivum cDNA clone wr1.pk0082.f4 5'
            end, mRNA sequence.
ACCESSION   CA607797
VERSION     CA607797.1  GI:25162959
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 265)
  AUTHORS   Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
  TITLE     DuPont Wheat cDNA Sequence
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
FEATURES             Location/Qualifiers
     source          1. .265
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone="wr1.pk0082.f4"
                     /tissue_type="root"
                     /clone_lib="wr1"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
                     XhoI; Wheat (Triticum aestivum L.) root; 7 day old
                     seedling, light grown"
ORIGIN

  Query Match             4.4%;  Score 98.8;  DB 14;  Length 265;
  Best Local Similarity   62.6%;  Pred. No. 2.4e-09;
  Matches  142;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy       1389 TAAACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGA 1448
             | ||     | |  |  ||  ||   || || ||||| |     ||  | |||  | ||  |
Db        230 TGAAGNTGACGANGNTGAAGAANAAGAAGAAGAAGAANAAGNAGATGAAGAAGAATA 171

Qy       1449 AGAAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGA 1508
             ||||   ||||| || |  | |||||| ||||  |  |||||||| |||       |||||
Db        170 AGAAGCAGAAGAAGANGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGANGAAGA 111

Qy       1509 GGATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGC 1568
             ||  |||  ||||| |||||||      |  |||||||||  | |||||||  || |||||
Db        110 AGANGATGANGAAGAAGAAGAAGANGANGAAGAAGAAGAAGCAGAAGAAGANGAAGAAGC 51

Qy       1569 CGAAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
             |||||  || ||||  ||||||   | ||||  ||      | |||||
Db         50 AGAAGAAGAAGAAGAAGAAGAAGAAGCAGAAGAAGAAGAAGAAGAAG 4


RESULT 9
AZ529191/c
LOCUS       AZ529191                880 bp    DNA     linear   GSS 03-NOV-2000
DEFINITION  ENTBV68TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, genomic survey sequence.
ACCESSION   AZ529191
VERSION     AZ529191.1  GI:11081835
```

KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 880)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica HM1:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 673.

FEATURES Location/Qualifiers
source 1. .880
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site_1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell, Oxford University Press, 1999)."

ORIGIN

Query Match 4.4%; Score 98.8; DB 28; Length 880;
Best Local Similarity 64.3%; Pred. No. 2.9e-09;
Matches 148; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```
Qy       1386 AAGTAAACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGA 1445
             ||   ||  | | |  | |||  |||  || || ||||| |  |  | ||  ||  ||||
Db        766 AATATAATTAGAAGACGCAGAAGAAGAAGACGAAGAAGATGAAGACGACGAAGAAGACGA 707

Qy       1446 CGAAGAAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGA 1505
             || |||   || || | || || ||| |||||||  | ||||||||||||        ||
Db        706 AGATGAAGAAGACGAAGCAGACGAAGATGACGAAGAAGACGAAGAAGACGAAGAAGAAGA 647

Qy       1506 AGAGGATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGA 1565
             |||  ||  | |  |||||||||||||| |   | |||||||||    || ||||  |||||
Db        646 AGACGAAGAAGAAGAAGACGAAGAAGAAGAAGACGAAGAAGATGAAGACGAAGAAGATGA 587

Qy       1566 AGCCGAAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
             ||  |||||| || ||||  ||||||  |||||  ||      | ||| |
Db        586 AGACGAAGAAGATGAAGACGAAGAAGATGAAGACGAAGAAGATGAAGACG 537
```

RESULT 10
BQ241566
LOCUS BQ241566 264 bp mRNA linear EST 03-MAY-2002
DEFINITION TaE05003E07R TaE05 Triticum aestivum cDNA clone TaE05003E07R, mRNA sequence.
ACCESSION BQ241566
VERSION BQ241566.1 GI:20437442
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 264)
AUTHORS Cloutier,S.
TITLE Wheat functional genomics - Glenlea developing seeds cDNA libraries
JOURNAL Unpublished (2002)
COMMENT Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.gc.ca
was cloned directionally, not all sequences generated with reverse primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >2.0 kb
Plate: 003 row: E column: 07
Seq primer: M13 Reverse.

FEATURES Location/Qualifiers
source 1. .264
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TaE05003E07R"
/tissue_type="developing seeds"
/dev_stage="5 days after anthesis"
/lab_host="E. coli DH10B"
/clone_lib="TaE05"
/note="Vector: pSPORT-P (Invitrogen Technologies); Site_1: NotI; Site_2: MluI; mRNA obtained from wheat seeds of cultivar Glenlea 5 days post-anthesis"

ORIGIN

Query Match 4.4%; Score 98.6; DB 13; Length 264;
Best Local Similarity 64.9%; Pred. No. 2.6e-09;
Matches 146; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy       1391 AACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAG 1450
             ||  | | | ||||||  |||  || || ||||| |  | ||  |  ||  | || ||||
Db         22 AAGAAGAAGTGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAAG 81

Qy       1451 AAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGG 1510
             ||    |||| ||||  | ||||| ||||  |  |||||||| |||       ||||| |
Db         82 AAGAAAAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 141

Qy       1511 ATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCG 1570
             |  | |  ||||| |||||||  |   | |||||||||   |||||||  || ||||  |
Db        142 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 201

Qy       1571 AAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
             |||| || ||||  ||||||   ||||||  ||       | |||||
Db        202 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 246
```

RESULT 11
CA566759
LOCUS CA566759 332 bp mRNA linear EST 19-NOV-2002
DEFINITION K0405E07-5N NIA Mouse Mesenchymal Stem Cell cDNA Library (Long) Mus musculus cDNA clone NIA:K0405E07 IMAGE:30059574 5', mRNA sequence.
ACCESSION CA566759
VERSION CA566759.1 GI:25111430
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 332)
AUTHORS Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G., Umezawa,A. and Ko,M.S.H.
TITLE Systematic Analyses of NIA Mouse Mesenchymal Stem Cell cDNA Library

(Long)
JOURNAL Unpublished (2001)
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0405 row: E column: 07
Seq primer: M13 Reverse
High quality sequence stop: 332
POLYA=No.

FEATURES Location/Qualifiers
source 1. .332
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H/He"
/db_xref="niaEST:K0405E07-5N"
/db_xref="taxon:10090"
/clone="NIA:K0405E07 IMAGE:30059574"
/tissue_type="Mesenchymal stem cell"
/cell_line="9-15-C cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library (Long)"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Akihiro Umezawa (Keio University School of Medicine, Japan). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from 2.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 4.4%; Score 98.6; DB 14; Length 332;
Best Local Similarity 64.9%; Pred. No. 2.7e-09;
Matches 146; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy       1391 AACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAG 1450
             ||  | | |  |||||  |||  || || ||||| |  | ||  |  ||  | || ||||
Db          3 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 62

Qy       1451 AAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGG 1510
             ||   ||||| ||||  | ||||||  ||||  |  ||||||| |||       |||||
Db         63 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 122

Qy       1511 ATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCG 1570
             |  | |  ||||| |||||||  |    |  |||||||| |||||||  || ||||  |
Db         123 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 182

Qy       1571 AAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
             ||||  ||  ||||  ||||||   |||||| ||       | |||||
Db         183 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 227
```

RESULT 12
BF707975
LOCUS BF707975 575 bp mRNA linear EST 12-JUN-2002

DEFINITION A634_LE Aplysia metacerebral cell cDNA library Aplysia californica cDNA 5', mRNA sequence.
ACCESSION BF707975
VERSION BF707975.1 GI:18002301
KEYWORDS EST.
SOURCE Aplysia californica (California sea hare)
ORGANISM Aplysia californica
Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda; Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidea; Aplysioidea; Aplysiidae; Aplysia.
REFERENCE 1 (bases 1 to 575)
AUTHORS Sadreyev,R.I., Meleshkevich,E.A., Matz,M.V., Panchin,Y.V. and Moroz,L.L.
TITLE Characterization of the gene pool expressed in the individual identified neuron
JOURNAL Unpublished (2001)
COMMENT Contact: Sadreyev RI
The Whitney Laboratory
University of Florida
9505 Ocean Shore Blvd., St Augustine, FL 32080, USA
Tel: 904 461 4044
Fax: 904 461 4008
Email: sadreyev@whitney.ufl.edu
Seq primer: 5' CGACGTGGACTATCCATGAACGCA 3'.

FEATURES Location/Qualifiers
source 1. .575
/organism="Aplysia californica"
/mol_type="mRNA"
/db_xref="taxon:6500"
/tissue_type="nervous system"
/cell_type="Metacerebral cell"
/clone_lib="Aplysia metacerebral cell cDNA library"
/note="Vector: pGEM-T"

ORIGIN

Query Match 4.4%; Score 98.6; DB 10; Length 575;
Best Local Similarity 64.9%; Pred. No. 3e-09;
Matches 146; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy       1391 AACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAG 1450
             ||  | | |  |||||  |||  || || ||||| |  | ||  |  ||  | || ||||
Db         105 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 164

Qy       1451 AAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGG 1510
             ||   ||||| ||||  | ||||||  ||||  |  ||||||| |||       |||||
Db         165 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 224

Qy       1511 ATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCG 1570
             |  | |  ||||| |||||||  |    |  |||||||| |||||||  || ||||  |
Db         225 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 284

Qy       1571 AAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAG 1615
             ||||  ||  ||||  ||||||   |||||| ||       | |||||
Db         285 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 329
```

RESULT 13
CC143364
LOCUS CC143364 960 bp DNA linear GSS 16-APR-2003
DEFINITION NDL.32A13.T7 Notre Dame Liverpool Aedes aegypti genomic clone NDL.32A13, genomic survey sequence.
ACCESSION CC143364
VERSION CC143364.1 GI:30012419
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes; Stegomyia.
REFERENCE 1 (bases 1 to 960)
AUTHORS Loftus,B., Shetty,J., Knudson,D. and Severson,D.
TITLE BAC end sequencing of Aedes aegypti
JOURNAL Unpublished (2003)

```
COMMENT     Other_GSSs: NDL.32A13.SP6
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: enta@tigr.org
            Library was provided by David Severson
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .960
                     /organism="Aedes aegypti"
                     /mol_type="genomic DNA"
                     /strain="liverpool"
                     /db_xref="taxon:7159"
                     /clone="NDL.32A13"
                     /clone_lib="Notre Dame Liverpool"
                     /note="Vector: pECBAC1; Site_1: Hind III; The library was
                     prepared from whole body tissue of newly hatched L1 larvae
                     by David Severson at the University of Notre Dame and
                     Hongbin Zhang"
ORIGIN

  Query Match             4.4%;  Score 98.4;  DB 28;  Length 960;
  Best Local Similarity   65.5%;  Pred. No. 3.5e-09;
  Matches  144;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy       1391 AACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAG 1450
             ||  | | |  ||||| ||||  || || ||||| |  | ||  |  ||  | || ||||
Db        438 AAGAAGAAGAAGAAGAAAAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 497

Qy       1451 AAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGG 1510
             ||   ||||| ||||  | |||||| ||||  |  |||||||| |||       ||||| |
Db        498 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 557

Qy       1511 ATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCG 1570
             |  | |  ||||| ||||||| |   |  |||||||||   |||||||  || ||||  |
Db        558 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 617

Qy       1571 AAGAGGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGC 1610
             |||| || ||||  ||||||   |||||| ||  |   ||
Db        618 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGTGGACGC 657


RESULT 14
BM167059
LOCUS       BM167059                 594 bp    mRNA    linear   EST 04-DEC-2001
DEFINITION  EST569582 PyBS Plasmodium yoelii yoelii cDNA clone PYCOC02 5' end,
            mRNA sequence.
ACCESSION   BM167059
VERSION     BM167059.1  GI:17300291
KEYWORDS    EST.
SOURCE      Plasmodium yoelii yoelii
  ORGANISM  Plasmodium yoelii yoelii
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 594)
  AUTHORS   Carlton,J.M., Daly,T.M., Long,C.A., Bergman,L.W., Vaidya,A.B.,
            Fraser,C.M. and Carucci,D.J.
  TITLE     Plasmodium yoelii EST project at TIGR
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Jane Carlton
            Parasite Genomics Group
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-530-9319
            Fax: 301-838-0208
            Email: carlton@tigr.org
            For clone info, please contact the Malaria Research and Reference
            Reagent Resource Center, ATCC
            http://www.malaria.mr4.org/mr4pages/index.html
            Seq primer: ADF.
FEATURES             Location/Qualifiers
     source          1. .594
                     /organism="Plasmodium yoelii yoelii"
                     /mol_type="mRNA"
                     /strain="17XL"
                     /sub_species="yoelii"
                     /db_xref="taxon:73239"
                     /clone="PYCOC02"
                     /dev_stage="Asexual blood stages"
                     /lab_host="E. coli XL-1 Blue"
                     /clone_lib="PyBS"
                     /note="Vector: pAD-GAL4; At 20-25% parasitemia, blood was
                     collected from BALB/cByJ mice infected with Py17XL
                     parasites, and leukocytes removed by passage over
                     microcrystalline cellulose collumns. Total RNA was
                     isolated using the guanidinium isothiocyanate method, and
                     mRNA isolated using oligo(dT)-cellulose chromatography.
                     First strand cDNA synthesis was completed using a 50-base
                     primer and reverse transcriptase in the presence of
                     5-methyl dCTP. After second strand synthesis, uneven
                     termini were treated with Pfu DNA polymerase and EcoRI
                     adaptors ligated to the blunt ends. The sample was cleaved
                     with XhoI and separated on a Sephacryl S-500 column.
                     Size-fractionated cDNA was precipitated and ligated to
                     HybriZAP arms directionally using EcoRI-XhoI cleaved arms.
                     After packaging, the phagemid vector (pAD-GAL4) was
                     excised from the HybriZAP vector and plasmid DNA
                     isolated."
ORIGIN

  Query Match             4.3%;  Score 98.2;  DB 12;  Length 594;
  Best Local Similarity   61.6%;  Pred. No. 3.6e-09;
  Matches  157;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy       1348 TATGTGAAACTCGGACGGATAAAAACCGATCGCCCGGCAAGTAAACCAAAGGCGGAAGAT 1407
             | | || |||| ||    |   |||   |||      | |     |  ||| | |||||
Db        285 TCTATGCAACTTGGGGAGCATAAATGGGATGAAGGGCAAGTGGACGAAAATGATGAAGAA 344

Qy       1408 AAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAGAAAGCGAAGACGAAGCC 1467
             ||||  ||||| ||||  |  | ||  |  ||  | || ||| ||  ||| || ||||
Db        345 AAAGAAGAGGAAGAAGTAGAAGAAGATGAAGAAGAGGAAGAAAAAGACGATGAGGAAGAG 404

Qy       1468 GTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGGATAATGGCGAAGACGAA 1527
             | |||||| |||   |  |||||||| |||| |   ||||| || ||||  ||||  |||
Db        405 GAAGAAGATGAAAATGAGGAAGAAGATGAAAATGAGGAAGATGAAAATGAGGAAGTAGAA 464

Qy       1528 GAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCGAAGAGGAGGAAGTTGAA 1587
             ||||      |   | ||||||   |||||  |  ||||||   ||||| |||||||  ||||
Db        465 GAAGATGATGAAAATGAAGAAGAGGAAGATGAAGATGAAAATGAAGAAGAGGAAGATGAA 524

Qy       1588 GAACCCGAAGAAAAA 1602
             ||    |||| | ||
Db        525 GATGAGGAAGTAGAA 539


RESULT 15
BZ149404
LOCUS       BZ149404                 612 bp    DNA     linear   GSS 11-OCT-2002
DEFINITION  CH230-395M1.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
            CH230-395M1, genomic survey sequence.
ACCESSION   BZ149404
VERSION     BZ149404.1  GI:23790354
KEYWORDS    GSS.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 612)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 MboI segment
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-395M1.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 395  row: M  column: 1
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .612
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-395M1"
                     /sex="Female"
                     /cell_type="Brain"
                     /clone_lib="CHORI-230 Segment 2"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
ORIGIN

  Query Match             4.3%;  Score 98.2;  DB 28;  Length 612;
  Best Local Similarity   64.1%;  Pred. No. 3.6e-09;
  Matches  148;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy       1395 AAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGCGTTGATAACGACGAAGAAAG 1454
             | ||| ||||||  ||| ||| || || || |  | ||  |  || || || ||||||
Db        157 AGAGGAGGAAGAGGAAGAGGAAGAGGAGGAAGAAGAAGAAGAAAAAGAAGAAGAAGA 216

Qy       1455 CGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGACGAAACTTCCGAAGAGGATAA 1514
              ||||| ||||  | |||||| ||||  |  |||||||| |||      || |||||  |
Db        217 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAGGATGAGGAGGA 276

Qy       1515 TGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAAGAAGTTGATGAAGCCGAAGA 1574
              |  || || || ||||      |  ||||| ||     |||||||  || ||||  |||||
Db        277 GGAGGAGGAAGAGGAAGAGGAAGAGGAAGAGGAGGAAGAAGAAGAAGAAGAAGAAGA 336

Qy       1575 GGAGGAAGTTGAAGAACCCGAAGAAAAATCGCCGGCAGAAGGCAACGGCGG 1625
              || ||||  ||||||   |||||| ||      | |||||  | |||||
Db        337 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGGCGG 387


Search completed: August 25, 2004, 05:21:12
Job time : 5163.56 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 08:53:30 ; Search time 7964.58 Seconds
(without alignments)
11558.757 Million cell updates/sec

Title: US-10-735-098-9
Perfect score: 2124
Sequence: 1 atgtgtaaaccgaattatgg.....acaaggaggcaacacgatga 2124

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		2124	100.0	2124	1	AF123383	AF123383 Neisseria
2		2124	100.0	2124	6	A98976	A98976 Sequence 9
3		2124	100.0	2124	6	BD074759	BD074759 Neisseria
4		1562.4	73.6	2169	1	AF123382	AF123382 Neisseria
5		1562.4	73.6	2169	6	A98970	A98970 Sequence 3
6		1562.4	73.6	2169	6	BD074756	BD074756 Neisseria
7	c	1516.8	71.4	326301	1	NMA6Z2491	AL162757 Neisseria
8		1499.8	70.6	2226	1	AF123380	AF123380 Neisseria
9		1499.8	70.6	2226	6	A98972	A98972 Sequence 5
10		1499.8	70.6	2226	6	BD074757	BD074757 Neisseria
11		1416	66.7	2519	1	AF072890	AF072890 Neisseria
12	c	1375.4	64.8	9955	1	AE002504	AE002504 Neisseria
13	c	1375.4	64.8	349980	6	AX044033	AX044033 Sequence
14		1296.2	61.0	2277	1	AF022781	AF022781 Neisseria
15		1296.2	61.0	2277	6	A98968	A98968 Sequence 1
16		1296.2	61.0	2277	6	BD074755	BD074755 Neisseria
17		1296.2	61.0	5691	1	AF049349	AF049349 Neisseria
18		1293	60.9	2537	1	AF031432	AF031432 Neisseria
19		1291.8	60.8	2262	1	AF123381	AF123381 Neisseria
20		1291.8	60.8	2262	6	A98974	A98974 Sequence 7
21		1291.8	60.8	2262	6	BD074758	BD074758 Neisseria
22		497.6	23.4	1000	6	AX081510	AX081510 Sequence
23		497.6	23.4	1000	6	AX374691	AX374691 Sequence
24		231.8	10.9	3398	1	NMLBPAG	X79838 N.meningiti
25	c	105.8	5.0	13297	10	AF317727	AF317727 Mus muscu
26		105.8	5.0	179685	10	AC122056	AC122056 Mus muscu
27		104.8	4.9	266939	2	AC098663	AC098663 Rattus no
28		104.6	4.9	3300	1	NGU16260	U16260 Neisseria g
29		104.6	4.9	3300	6	AX685922	AX685922 Sequence
30	c	104	4.9	224784	2	AC110522	AC110522 Mus muscu
31		103.6	4.9	196963	2	AC133659	AC133659 Rattus no
32	c	103.4	4.9	173949	10	AC132100	AC132100 Mus muscu
33	c	103.4	4.9	187930	2	AC124191	AC124191 Mus muscu
34		103.2	4.9	194060	2	AC125897	AC125897 Rattus no
35		103	4.8	163710	2	AC112834	AC112834 Rattus no
36		103	4.8	177128	2	AC113041	AC113041 Mus muscu
37		103	4.8	222469	10	AC140386	AC140386 Mus muscu
38	c	103	4.8	224843	2	AC147135	AC147135 Mus muscu
39		103	4.8	240033	2	AC133684	AC133684 Rattus no
40	c	103	4.8	251725	2	AC094726	AC094726 Rattus no
41		102.8	4.8	171477	2	AC102467	AC102467 Mus muscu
42	c	102.6	4.8	173556	2	AC074388	AC074388 Homo sapi
43	c	102.4	4.8	123180	10	AL672005	AL672005 Mouse DNA
44		102.4	4.8	232224	2	AC110149	AC110149 Rattus no
45	c	102.4	4.8	251960	2	AC107279	AC107279 Rattus no

ALIGNMENTS

RESULT 1
AF123383
LOCUS AF123383 2124 bp DNA linear BCT 24-MAY-1999
DEFINITION Neisseria meningitidis strain 881607 lactoferrin-binding protein precursor (lbpB) gene, complete cds.
ACCESSION AF123383
VERSION AF123383.1 GI:4884692
KEYWORDS .
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 2124)
AUTHORS Pettersson,A., van der Biezen,J., Joosten,V., Hendriksen,J. and Tommassen,J.

```
  TITLE     Sequence variability of the meningococcal lactoferrin-binding
            protein LbpB
  JOURNAL   Gene 231 (1-2), 105-110 (1999)
  MEDLINE   99250255
   PUBMED   10231574
REFERENCE   2  (bases 1 to 2124)
  AUTHORS   Pettersson,A., van der Biezen,J., Jooster,V., Hendriksen,J. and
            Tommassen,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-1999) Department of Molecular Cell Biology,
            Utrecht University, Padualaan 8, Utrecht 3584 CH, The Netherlands
FEATURES             Location/Qualifiers
     source          1. .2124
                     /organism="Neisseria meningitidis"
                     /mol_type="genomic DNA"
                     /strain="881607"
                     /db_xref="taxon:487"
     gene            1. .2124
                     /gene="lbpB"
     CDS             1. .2124
                     /gene="lbpB"
                     /codon_start=1
                     /transl_table=11
                     /product="lactoferrin-binding protein precursor"
                     /protein_id="AAD31771.1"
                     /db_xref="GI:4884693"
                     /translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
                     DVPTSPPAGSSVETTPVNRPAVGAAMRLLRRNIATSDKDGNDFPNSKQAEEKLSFKEE
                     DILFLYGSKKDQRQQLKDKIRQPNPTASITTSEKKNKKYDYKFVDAGYVYTKDGKDEI
                     EWTSNYKQSTNRFGYDGFVYYSGEHPSQSLPSAGTVKYSGNWQYMTDAIRHRTGKAGD
                     PSEDLGYIVYYGQNVGATSYAATADDREGKHPAEYTVNFDQKTLNGKLIKNQYVQKRD
                     DPKKPLTIYDITAKLDGNRFTGSAKVNTEVKTNHADKEYLFFHTDADQRLEGGFFGDK
                     GEELAGRFISNDNSVFGVFAGKQKTETANASDTNPALPSGKHTKILDSLKISVDEASG
                     ENPRPFEVSTMPDFGHPDKLLVEGREIPLVNKEQTIDLADGRKMTVRACCDFLTYVKL
                     GRIKTERPAVQPKAQDEEGDEEGVGVDNGKESEDEIGDEESTGDEVVEDEDEDEDEEE
                     IEEEPEEEAEEEEPEEELPAEEGNGGSGSILPTPEASKGRDIDLFLKGIRTAEADIPK
                     NGTAHYTGTWEARIGVSDSGTSIQKDSYANQGAKAEFTVDFEAKTVSGMLTEKNDTTP
                     AFYIEKGVIDGNGFHALAHTRENGIDLSGQGSTNPKNFKADNLLVTGGFYGPQAAELG
                     GNIIDSDRKFGAVFGAKKDDKEATR"
ORIGIN

  Query Match             100.0%;  Score 2124;  DB 1;  Length 2124;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2124;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180

Qy        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240

Qy        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300

Qy        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360

Qy        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420

Qy        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480

Qy        481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540

Qy        541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600

Qy        601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660

Qy        661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720

Qy        721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780

Qy        781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840

Qy        841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900

Qy        901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960

Qy        961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020

Qy       1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080

Qy       1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140

Qy       1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200

Qy       1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260

Qy       1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320

Qy       1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380

Qy       1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440

Qy       1441 GAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAA 1500

Qy       1501 GACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAA 1560
```

```
Db       1501 GACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAA 1560

Qy       1561 GAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCC 1620
Db       1561 GAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCC 1620

Qy       1621 ACTCCGGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCG 1680
Db       1621 ACTCCGGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCG 1680

Qy       1681 GAAGCCGACATTCCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATC 1740
Db       1681 GAAGCCGACATTCCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATC 1740

Qy       1741 GGCGTATCGGATAGTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAA 1800
Db       1741 GGCGTATCGGATAGTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAA 1800

Qy       1801 GCAGAATTTACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAAT 1860
Db       1801 GCAGAATTTACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAAT 1860

Qy       1861 GATACAACCCCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCT 1920
Db       1861 GATACAACCCCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCT 1920

Qy       1921 TTGGCGCATACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAG 1980
Db       1921 TTGGCGCATACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAG 1980

Qy       1981 AACTTCAAAGCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAA 2040
Db       1981 AACTTCAAAGCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAA 2040

Qy       2041 TTGGGCGGTAATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAA 2100
Db       2041 TTGGGCGGTAATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAA 2100

Qy       2101 GATGACAAGGAGGCAACACGATGA 2124
Db       2101 GATGACAAGGAGGCAACACGATGA 2124
```

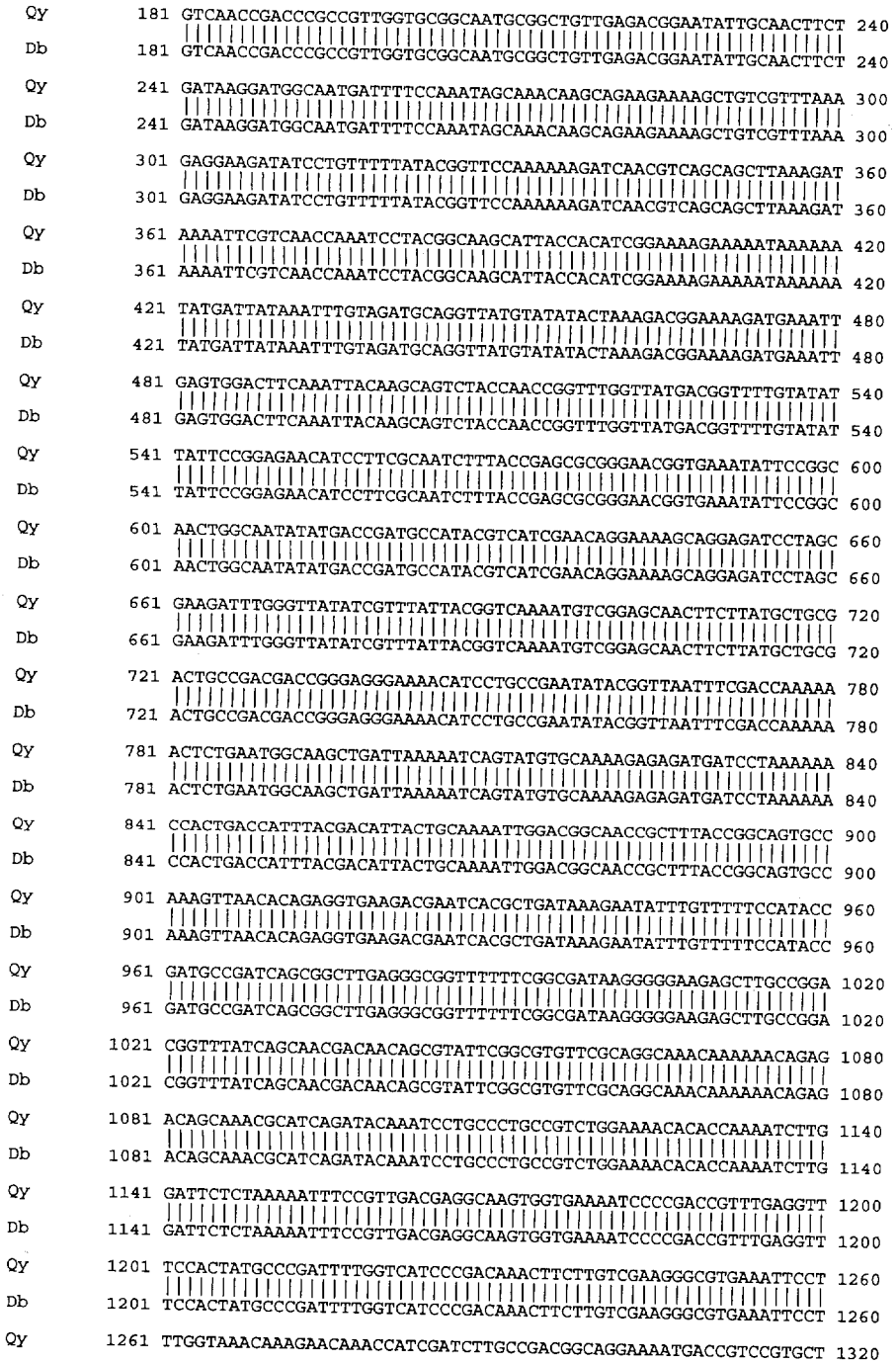
```
RESULT 2
A98976
LOCUS       A98976                2124 bp    DNA     linear   PAT 26-JAN-2000
DEFINITION  Sequence 9 from Patent WO9909176.
ACCESSION   A98976
VERSION     A98976.1  GI:6781936
KEYWORDS    .
SOURCE      Neisseria meningitidis
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE   1  (bases 1 to 2124)
  AUTHORS   Pettersson-Fernholm,A.M. and Tommassen,J.P.
  TITLE     NEISSERIA LACTOFERRIN BINDING PROTEIN
  JOURNAL   Patent: WO 9909176-A 9 25-FEB-1999;
            UNIV UTRECHT (NL); PETTERSSON FERNHOLM ANNIKA MAR (NL)
FEATURES             Location/Qualifiers
     source          1. .2124
                     /organism="Neisseria meningitidis"
                     /mol_type="unassigned DNA"
                     /strain="881607"
                     /db_xref="taxon:487"
     CDS             1. .2124
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB70585.1"
                     /db_xref="GI:6781937"
                     /db_xref="REMTREMBL:CAB70585"
                     /translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
                     DVPTSPPAGSSVETTPVNRPAVGAAMRLLRRNIATSDKDGNDFPNSKQAEEKLSFKEE
                     DILFLYGSKKDQRQQLKDKIRQPNPTASITTSEKKNKKYDYKFVDAGYVYTKDGKDEI
                     EWTSNYKQSTNRFGYDGFVYYSGEHPSQSLPSAGTVKYSGNWQYMTDAIRHRTGKAGD
                     PSEDLGYIVYYGQNVGATSYAATADDREGKHPAEYTVNFDQKTLNGKLIKNQYVQKRD
                     DPKKPLTIYDITAKLDGNRFTGSAKVNTEVKTNHADKEYLFFHTDADQRLEGGFFGDK
                     GEELAGRFISNDNSVFGVFAGKQKTETANASDTNPALPSGKHTKILDSLKISVDEASG
                     ENPRPFEVSTMPDFGHPDKLLVEGREIPLVNKEQTIDLADGRKMTVRACCDFLTYVKL
                     GRIKTERPAVQPKAQDEEGDEEGVGVDNGKESEDEIGDEESTGDEVVEDEDEDEDEEE
                     IEEEPEEEAEEEEPEEELPAEEGNGGSGSILPTPEASKGRDIDLFLKGIRTAEADIPK
                     NGTAHYTGTWEARIGVSDSGTSIQKDSYANQGAKAEFTVDFEAKTVSGMLTEKNDTTP
                     AFYIEKGVIDGNGFHALAHTRENGIDLSGQGSTNPKNFKADNLLVTGGFYGPQAAELG
                     GNIIDSDRKFGAVFGAKKDDKEATR"
ORIGIN

  Query Match           100.0%;  Score 2124;  DB 6;  Length 2124;
  Best Local Similarity 100.0%;  Pred. No. 0;
  Matches 2124;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180

Qy        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
Db        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240

Qy        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
Db        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300

Qy        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
Db        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360

Qy        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
Db        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420

Qy        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
Db        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480

Qy        481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
Db        481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540

Qy        541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
Db        541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600

Qy        601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
Db        601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660

Qy        661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
Db        661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720

Qy        721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
Db        721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780

Qy        781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
```

```
Db        781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
Qy        841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
Db        841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
Qy        901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
Db        901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
Qy        961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
Db        961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
Qy       1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
Db       1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
Qy       1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
Db       1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
Qy       1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
Db       1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
Qy       1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
Db       1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
Qy       1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
Db       1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
Qy       1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
Db       1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
Qy       1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
Db       1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
Qy       1441 GAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAA 1500
Db       1441 GAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAA 1500
Qy       1501 GACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAA 1560
Db       1501 GACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAA 1560
Qy       1561 GAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCC 1620
Db       1561 GAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCC 1620
Qy       1621 ACTCCGGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCG 1680
Db       1621 ACTCCGGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCG 1680
Qy       1681 GAAGCCGACATTCCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATC 1740
Db       1681 GAAGCCGACATTCCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATC 1740
Qy       1741 GGCGTATCGGATAGTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAA 1800
Db       1741 GGCGTATCGGATAGTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAA 1800
Qy       1801 GCAGAATTTACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAAT 1860
Db       1801 GCAGAATTTACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAAT 1860
Qy       1861 GATACAACCCCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCT 1920
Db       1861 GATACAACCCCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCT 1920
Qy       1921 TTGGCGCATACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAG 1980
Db       1921 TTGGCGCATACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAG 1980
Qy       1981 AACTTCAAAGCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAA 2040
Db       1981 AACTTCAAAGCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAA 2040
Qy       2041 TTGGGCGGTAATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAA 2100
Db       2041 TTGGGCGGTAATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAA 2100
Qy       2101 GATGACAAGGAGGCAACACGATGA 2124
Db       2101 GATGACAAGGAGGCAACACGATGA 2124
```

RESULT 3
BD074759
LOCUS BD074759 2124 bp DNA linear PAT 27-AUG-2002
DEFINITION Neisseria lactoferrin-binding protein.
ACCESSION BD074759
VERSION BD074759.1 GI:22620362
KEYWORDS JP 2001514894-A/5.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 2124)
AUTHORS Felnholm,A.M.P. and Thomasen,J.P.M.
TITLE Neisseria lactoferrin-binding protein
JOURNAL Patent: JP 2001514894-A 5 18-SEP-2001;
UNIVERSITY OF UTRECHT,TECHNOLOGY FOUNDATION
COMMENT OS Unidentified
PN JP 2001514894-A/5
PD 18-SEP-2001
PF 10-AUG-1998 JP 2000509840
PR 15-AUG-1997 GB 9717423.9,05-FEB-1998 GB 9802544.8 PI
ANICA MARGARETA PETERSON FELNHOLM,JOHANES PETRUS MARIA PI
THOMASEN
PC C12N15/09,A61K39/095,A61K39/395,A61K48/00,A61P31/12,C07K14/22,
PC C07K16/12,
PC C12N1/21,C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/68, PC
C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
CC Neisseria lactoferrin-binding protein
FH Key Location/Qualifiers
FT source 1. .2124
FT /organism='Unidentified'.
FEATURES Location/Qualifiers
source 1. .2124
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 100.0%; Score 2124; DB 6; Length 2124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
```



RESULT 4
AF123382
LOCUS AF123382 2169 bp DNA linear BCT 24-MAY-1999
DEFINITION Neisseria meningitidis strain m981 lactoferrin-binding protein precursor (lbpB) gene, complete cds.
ACCESSION AF123382
VERSION AF123382.1 GI:4884690
KEYWORDS .
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 2169)

```
  AUTHORS   Pettersson,A., van der Biezen,J., Joosten,V., Hendriksen,J. and
            Tommassen,J.
  TITLE     Sequence variability of the meningococcal lactoferrin-binding
            protein LbpB
  JOURNAL   Gene 231 (1-2), 105-110 (1999)
  MEDLINE   99250255
   PUBMED   10231574
REFERENCE   2  (bases 1 to 2169)
  AUTHORS   Pettersson,A., van der Biezen,J., Joosten,V., Hendriksen,J. and
            Tommassen,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-1999) Department of Molecular Cell Biology,
            Utrecht University, Padualaan 8, Utrecht 3584 CH, The Netherlands
FEATURES             Location/Qualifiers
     source          1. .2169
                     /organism="Neisseria meningitidis"
                     /mol_type="genomic DNA"
                     /strain="m981"
                     /db_xref="taxon:487"
     gene            1. .2169
                     /gene="lbpB"
     CDS             1. .2169
                     /gene="lbpB"
                     /codon_start=1
                     /transl_table=11
                     /product="lactoferrin-binding protein precursor"
                     /protein_id="AAD31770.1"
                     /db_xref="GI:4884691"
                     /translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
                     DVPTSPPAGSSVETTPVNQPAVGAAMRLLRRNTAFHREDGTAIPDSKQAEEKLSFKEG
                     DVLFLYGSKENKLQQLKSEIHKRNPEASITTSENENKKYNYRFVSAGYVFTKNGKDEI
                     EKTSDEKQFSNRLGYDGFVYYLGEHPSQSLPSAGTVKYSGNWQYMTDAIRHRRGKGVS
                     SVDLGYTTYYGNEIGAASYEARDADGREKHPAEYTVNFDKKNLEGKLIKNQYVQKRDD
                     PKNPLTIYNITATLDGNRFTGSAKVSTEVKTQHADKEYLFFHTDADQRLEGGFFGDNG
                     EELAGRFISNDNSVFGVFAGKQKTETANASDTNPALPSGKHTKILDSLKISVDEATDD
                     HARKFAISTMPDFGHPDKLLVEGREIPLVSQEKTIELADGRKMTIRACCDFLTYVKLG
                     RIKTDRPAVKPKAQDEEDSDIDNGEESEDEISEDDNGEDEVTEEEEAEETEEETDEDE
                     EEEPEETEETEETEETEETEETEEKSPTEEGNGGSGSILPTPEASKGRDIDLFLKGIR
                     TAEADIPQIGKARYTGTWEARIGVPDKKGEQLDGTTSIQKDSYANQAAKAEFDVDFGA
                     KSLSGKLTEKNDTHPAFYIEKGVIDGNGFHALARTRENGVDLSGQGSTNPQSFKASNL
                     LVEGGFYGPQAAELGGNIIDSDRKIGVVFGAKKDMQEVEK"
ORIGIN

  Query Match             73.6%;  Score 1562.4;  DB 1;  Length 2169;
  Best Local Similarity   84.5%;  Pred. No. 0;
  Matches 1835;  Conservative    0;  Mismatches  286;  Indels   51;  Gaps    5;

Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCCCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180

Qy        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
              |||||||  |||||||| |||||||||||||||||||||||||||| |||    | |
Db        181 GTCAACCAGCCCGCCGTCGGTGCGGCAATGCGGCTGTTGAGACGGAATACTGCTTTTCAT 240

Qy        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
                \ \ \\\\\\\  |   |||| ||||||||||||||||||||||||||||||||||||
Db        241 CGTGAAGATGGCACGGCAATTCCCGATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300

Qy        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
              || |  ||| | |||||||||||||||||| ||| || || ||| ||| || ||||||
Db        301 GAAGGTGATGTTCTGTTTTTATACGGTTCAAAAGAAAATAAACTTCAACAACTTAAAAGC 360

Qy        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
              ||||||  | |||  ||||||  |||||||||||||||||||||||  ||||||||||

Db        361 GAAATTCATAAACGTAATCCTGAGGCAAGCATTACCACATCGGAAAATGAAAATAAAAAA 420

Qy        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
              ||| |||||   |||||   ||| |||||||| | |||||||  |||||||||||||||||
Db        421 TATAATTATCGGTTTGTCAGTGCCGGTTATGTGTTTACTAAAAACGGAAAAGATGAAATT 480

Qy        481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
              |||   || ||  || | ||||||| | | || || || || ||||||||||||||||||
Db        481 GAGAAAACATCGGATGAAAAGCAGTTTTCTAATCGTTTAGGCTATGACGGTTTTGTATAT 540

Qy        541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
              |||  ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        541 TATCTCGGAGAACATCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600

Qy        601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
              |||||||||||||||||||||||||||||||||||||| ||| ||  || | | | ||
Db        601 AACTGGCAATATATGACCGATGCCATACGTCATCGGAGAGGTAA---GGGGGTTTCCAGT 657

Qy        661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
              | ||||||||||||| | |||| ||| |  |  | || ||| |||||||| ||
Db        658 GTGGATTTGGGTTATACCACATATTATGGTAATGAAATTGGGGCAGCTTCTTATGAGGCT 717

Qy        721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
              | | | |||  |  || ||||||||||||||||||||||||||||||||| |||||
Db        718 AGGGATGCCGATGGCCGGGAAAAACATCCTGCCGAATATACGGTTAATTTCGACAAAAAA 777

Qy        781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
              | ||| | || ||| ||||||||||||||||||||||||||||||||||||||||||||
Db        778 AACCTGGAAGGTAAGTTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAT 837

Qy        841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
              |||||||||||||| ||||||| |||| ||||||||||||||||||||||||||||||
Db        838 CCACTGACCATTTACAACATTACCGCAACATTGGACGGCAACCGCTTTACCGGCAGTGCC 897

Qy        901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
              ||||||| ||| |||||||||||| | ||||||||||||||||||||||||||||||||
Db        898 AAAGTTAGCACCGAGGTGAAGACGCAACACGCTGATAAAGAATATTTGTTTTTCCATACC 957

Qy        961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
              |||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db        958 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAACGGAGAAGAGCTTGCCGGG 1017

Qy       1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 CGGTTTATCAGTAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1077

Qy       1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1137

Qy       1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
              |||||||||||||||||||||||||||||||| | || ||| || |||| |||||  ||
Db       1138 GATTCTCTAAAAATTTCCGTTGACGAGGCGACTGATGACCATGCCCGTAAGTTTGCCATT 1197

Qy       1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1257

Qy       1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
              |||||  | ||||  |||||||||| ||||||||||||||||||||||| ||||||||
Db       1258 TTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCATCCGTGCT 1317

Qy       1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
              |||||||| ||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 TGTTGCGATTTTCTGACCTATGTGAAACTCGGACGGATAAAAACCGACCGCCCCGCCGTC 1377

Qy       1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
              |||||||||||||||||||||||||  ||   | |  |||| || | ||   |
Db       1378 AAACCGAAGGCGCAGGATGAAGAGGATTCGGACATTGATAATGGCGAAGAAAGCGAAGAC 1437

Qy       1441 GAAA---GCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGA---------------- 1481
              ||||    |||||| || | ||||||| || |||  ||||| |||
Db       1438 GAAATTTCCGAAGATGATAACGGCGAAGATGAAGTCACCGAAGAAGAGGAAGCTGAAGAA 1497
```

```
Qy       1482 --CGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAA------GAA 1533
Db       1498 ACCGAAGAAGAAACTGATGAAGACGAAGAGGAAGAACCCGAAGAAACTGAAGAAACTGAA 1557

Qy       1534 GAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGC 1593
Db       1558 GAAACTGAAGAAACTGAAGAAACTGAAGAAAAATCGCCGACAGAAGAAGGC 1617

Qy       1594 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1653
Db       1618 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1677

Qy       1654 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCAAAAAACGGAACGGCGCAT 1713
Db       1678 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCGCAAATTGGAAAAGCACGC 1737

Qy       1714 TATACCGGCACTTGGGAAGCGCGTATCGGCGTATCGGATA-------------------- 1753
Db       1738 TATACCGGCACTTGGGAAGCGCGTATCGGCGTGCCGGATAAGAAAGGCGAACAGCTAGAT 1797

Qy       1754 -GTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAAGCAGAATTTACC 1812
Db       1798 GGCACTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGCGGCAAAAGCAGAATTTGAC 1857

Qy       1813 GTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAATGATACAACCCCC 1872
Db       1858 GTTGATTTTGGTGCGAAGTCGCTTTCAGGTAAGTTGACAGAAAAAAATGATACACACCCC 1917

Qy       1873 GCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCTTTGGCGCATACT 1932
Db       1918 GCTTTTTATATTGAAAAAGGTGTGATTGATGGCAACGGTTTCCACGCTTTGGCGCGTACT 1977

Qy       1933 CGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAAGCC 1992
Db       1978 CGTGAAAATGGTGTTGATTTGTCTGGGCAAGGTTCGACTAATCCCCAAAGTTTTAAAGCC 2037

Qy       1993 GACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGTAAT 2052
Db       2038 AGTAATCTTCTCGTAGAAGGAGGATTTTATGGTCCGCAGGCGGCAGAGTTGGGTGGTAAT 2097

Qy       2053 ATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGACAAGGAG 2112
Db       2098 ATTATCGACAGTGACCGGAAAATCGGCGTGGTATTCGGTGCGAAGAAAGATATGCAGGAG 2157

Qy       2113 GCAACACGATGA 2124
Db       2158 GTGGAAAAATGA 2169
```

RESULT 5
A98970

```
LOCUS       A98970                 2169 bp    DNA     linear   PAT 26-JAN-2000
DEFINITION  Sequence 3 from Patent WO9909176.
ACCESSION   A98970
VERSION     A98970.1  GI:6781930
KEYWORDS    .
SOURCE      Neisseria meningitidis
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE   1  (bases 1 to 2169)
  AUTHORS   Pettersson-Fernholm,A.M. and Tommassen,J.P.
  TITLE     NEISSERIA LACTOFERRIN BINDING PROTEIN
  JOURNAL   Patent: WO 9909176-A 3 25-FEB-1999;
            UNIV UTRECHT (NL); PETTERSSON FERNHOLM ANNIKA MAR (NL)
FEATURES             Location/Qualifiers
     source          1. .2169
                     /organism="Neisseria meningitidis"
                     /mol_type="unassigned DNA"
                     /strain="M981"
                     /db_xref="taxon:487"
     CDS             1. .2169
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB70582.1"
                     /db_xref="GI:6781931"
                     /db_xref="REMTREMBL:CAB70582"
                     /translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
                     DVPTSPPAGSSVETTPVNQPAVGAAMRLLRRNTAFHREDGTAIPDSKQAEEKLSFKEG
                     DVLFLYGSKENKLQQLKSEIHKRNPEASITTSENENKKYNYRFVSAGYVFTKNGKDEI
                     EKTSDEKQFSNRLGYDGFVYYLGEHPSQSLPSAGTVKYSGNWQYMTDAIRHRRGKGVS
                     SVDLGYTTYYGNEIGAASYEARDADGREKHPAEYTVNFDKKNLEGKLIKNQYVQKRDD
                     PKNPLTIYNITATLDGNRFTGSAKVSTEVKTQHADKEYLFFHTDADQRLEGGFFGDNG
                     EELAGRFISNDNSVFGVFAGKQKTETANASDTNPALPSGKHTKILDSLKISVDEATDD
                     HARKFAISTMPDFGHPDKLLVEGREIPLVSQEKTIELADGRKMTIRACCDFLTYVKLG
                     RIKTDRPAVKPKAQDEEDSDIDNGEESEDEISEDDNGEDEVTEEEEAEETEEETDEDE
                     EEEPEETEETEETEETEETEETEEKSPTEEGNGGSGSILPTPEASKGRDIDLFLKGIR
                     TAEADIPQIGKARYTGTWEARIGVPDKKGEQLDGTTSIQKDSYANQAAKAEFDVDFGA
                     KSLSGKLTEKNDTHPAFYIEKGVIDGNGFHALARTRENGVDLSGQGSTNPQSFKASNL
                     LVEGGFYGPQAAELGGNIIDSDRKIGVVFGAKKDMQEVEK"
ORIGIN
```

Query Match 73.6%; Score 1562.4; DB 6; Length 2169;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1835; Conservative 0; Mismatches 286; Indels 51; Gaps 5;

```
Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCCCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180

Qy        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
Db        181 GTCAACCAGCCCGCCGTCGGTGCGGCAATGCGGCTGTTGAGACGGAATACTGCTTTTCAT 240

Qy        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
Db        241 CGTGAAGATGGCACGGCAATTCCCGATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300

Qy        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
Db        301 GAAGGTGATGTTCTGTTTTTATACGGTTCAAAAGAAAATAAACTTCAACAACTTAAAAGC 360

Qy        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
Db        361 GAAATTCATAAACGTAATCCTGAGGCAAGCATTACCACATCGGAAAATGAAAATAAAAAA 420

Qy        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
Db        421 TATAATTATCGGTTTGTCAGTGCCGGTTATGTGTTTACTAAAAACGGAAAAGATGAAATT 480

Qy        481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
Db        481 GAGAAAACATCGGATGAAAAGCAGTTTTCTAATCGTTTAGGCTATGACGGTTTTGTATAT 540

Qy        541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
Db        541 TATCTCGGAGAACATCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600

Qy        601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
Db        601 AACTGGCAATATATGACCGATGCCATACGTCATCGGAGAGGTAA---GGGGGTTTCCAGT 657

Qy        661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
Db        658 GTGGATTTGGGTTATACCACATATTATGGTAATGAAATTGGGGCAGCTTCTTATGAGGCT 717
```

```
Qy       721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
Db       718 AGGGATGCCGATGGCCGGGAAAAACATCCTGCCGAATATACGGTTAATTTCGACAAAAAA 777

Qy       781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
Db       778 AACCTGGAAGGTAAGTTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAT 837

Qy       841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
Db       838 CCACTGACCATTTACAACATTACCGCAACATTGGACGGCAACCGCTTTACCGGCAGTGCC 897

Qy       901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
Db       898 AAAGTTAGCACCGAGGTGAAGACGCAACACGCTGATAAAGAATATTTGTTTTTCCATACC 957

Qy       961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
Db       958 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAACGGAGAAGAGCTTGCCGGG 1017

Qy      1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
Db      1018 CGGTTTATCAGTAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1077

Qy      1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
Db      1078 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1137

Qy      1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
Db      1138 GATTCTCTAAAAATTTCCGTTGACGAGGCGACTGATGACCATGCCCGTAAGTTTGCCATT 1197

Qy      1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
Db      1198 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1257

Qy      1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
Db      1258 TTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCATCCGTGCT 1317

Qy      1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
Db      1318 TGTTGCGATTTTCTGACCTATGTGAAACTCGGACGGATAAAAACCGACCGCCCCGCCGTC 1377

Qy      1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
Db      1378 AAACCGAAGGCGCAGGATGAAGAGGATTCGGACATTGATAATGGCGAAGAAAGCGAAGAC 1437

Qy      1441 GAAA---GCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGA---------------- 1481
Db      1438 GAAATTTCCGAAGATGATAACGGCGAAGATGAAGTCACCGAAGAAGAGGAAGCTGAAGAA 1497

Qy      1482 --CGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAA------GAA 1533
Db      1498 ACCGAAGAAGAAACTGATGAAGACGAAGAGGAAGAACCCGAAGAAACTGAAGAAACTGAA 1557

Qy      1534 GAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGC 1593
Db      1558 GAAACTGAAGAAACTGAAGAAACTGAAGAAACTGAAGAAAAATCGCCGACAGAAGAAGGC 1617

Qy      1594 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1653
Db      1618 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1677

Qy      1654 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCAAAAAACGGAACGGCGCAT 1713
Db      1678 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCGCAAATTGGAAAAGCACGC 1737

Qy      1714 TATACCGGCACTTGGGAAGCGCGTATCGGCGTATCGGATA-------------------- 1753
Db      1738 TATACCGGCACTTGGGAAGCGCGTATCGGCGTGCCGGATAAGAAAGGCGAACAGCTAGAT 1797

Qy      1754 -GTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAAGCAGAATTTACC 1812

Db      1798 GGCACTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGCGGCAAAAGCAGAATTTGAC 1857

Qy      1813 GTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAATGATACAACCCCC 1872
Db      1858 GTTGATTTTGGTGCGAAGTCGCTTTCAGGTAAGTTGACAGAAAAAAATGATACACACCCC 1917

Qy      1873 GCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCTTTGGCGCATACT 1932
Db      1918 GCTTTTTATATTGAAAAAGGTGTGATTGATGGCAACGGTTTCCACGCTTTGGCGCGTACT 1977

Qy      1933 CGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAAGCC 1992
Db      1978 CGTGAAAATGGTGTTGATTTGTCTGGGCAAGGTTCGACTAATCCCCAAAGTTTTAAAGCC 2037

Qy      1993 GACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGTAAT 2052
Db      2038 AGTAATCTTCTCGTAGAAGGAGGATTTTATGGTCCGCAGGCGGCAGAGTTGGGTGGTAAT 2097

Qy      2053 ATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGACAAGGAG 2112
Db      2098 ATTATCGACAGTGACCGGAAAATCGGCGTGGTATTCGGTGCGAAGAAAGATATGCAGGAG 2157

Qy      2113 GCAACACGATGA 2124
Db      2158 GTGGAAAAATGA 2169
```

RESULT 6
BD074756

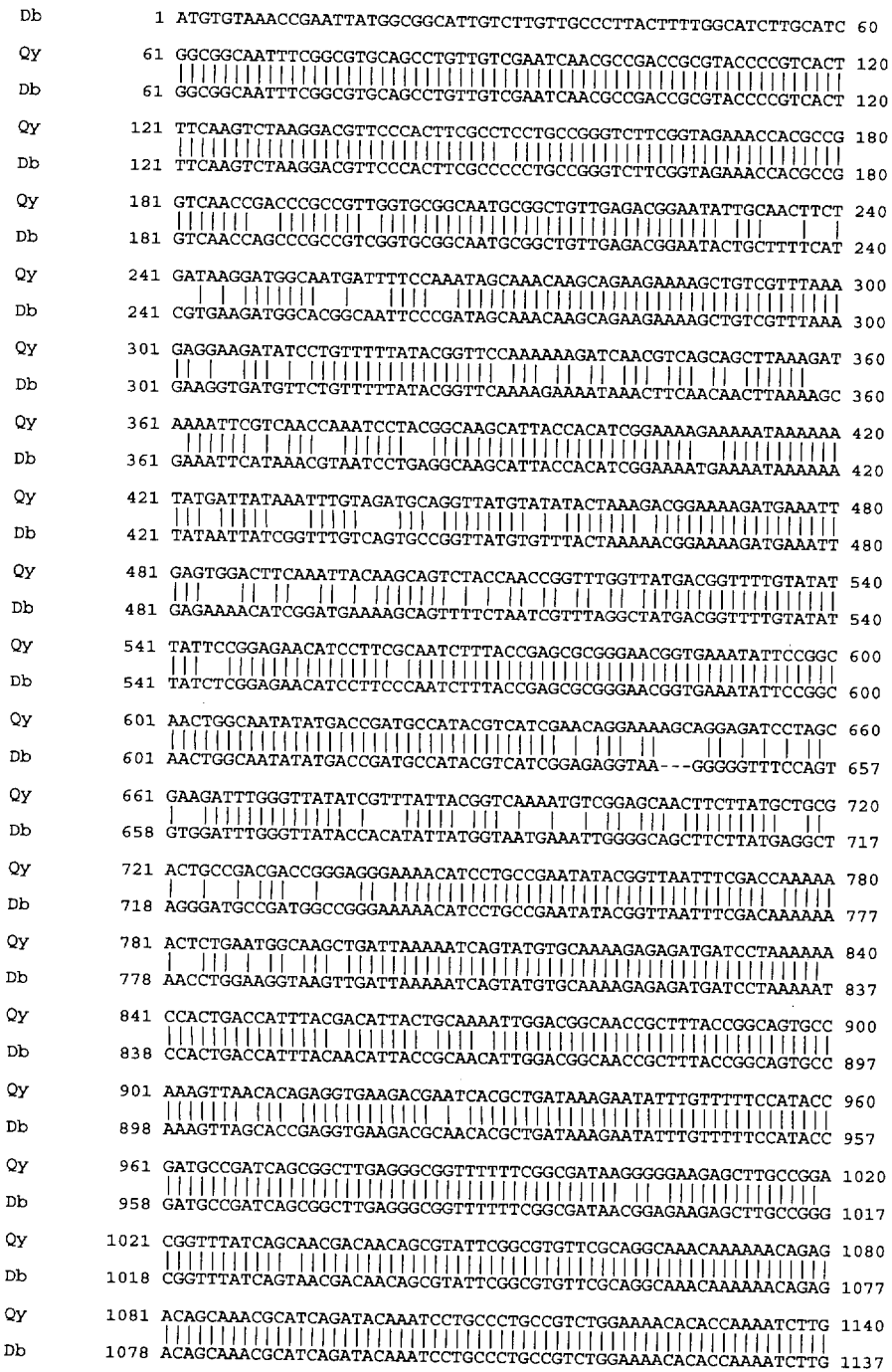
LOCUS	BD074756	2169 bp DNA linear PAT 27-AUG-2002
DEFINITION	Neisseria lactoferrin-binding protein.	
ACCESSION	BD074756	
VERSION	BD074756.1 GI:22620359	
KEYWORDS	JP 2001514894-A/2.	
SOURCE	unidentified	
ORGANISM	unidentified	
	unclassified.	
REFERENCE	1 (bases 1 to 2169)	
AUTHORS	Felnholm,A.M.P. and Thomasen,J.P.M.	
TITLE	Neisseria lactoferrin-binding protein	
JOURNAL	Patent: JP 2001514894-A 2 18-SEP-2001;	
	UNIVERSITY OF UTRECHT,TECHNOLOGY FOUNDATION	
COMMENT	OS Unidentified	
	PN JP 2001514894-A/2	
	PD 18-SEP-2001	
	PF 10-AUG-1998 JP 2000509840	
	PR 15-AUG-1997 GB 9717423.9,05-FEB-1998 GB 9802544.8 PI	
	ANICA MARGARETA PETERSON FELNHOLM,JOHANES PETRUS MARIA PI	
	THOMASEN	
	PC C12N15/09,A61K39/095,A61K39/395,A61K48/00,A61P31/12,C07K14/22,	
	PC C07K16/12,	
	PC C12N1/21,C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/68, PC	
	C12N15/00	
	CC Strandedness: Double;	
	CC Topology: Linear;	
	CC Neisseria lactoferrin-binding protein	
	FH Key Location/Qualifiers	
	FT source 1. .2169	
	FT /organism='Unidentified'.	

FEATURES	Location/Qualifiers
source	1. .2169
	/organism="unidentified"
	/mol_type="genomic DNA"
	/db_xref="taxon:32644"

ORIGIN

Query Match 73.6%; Score 1562.4; DB 6; Length 2169;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1835; Conservative 0; Mismatches 286; Indels 51; Gaps 5;

```
Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
```


Db 1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Qy 61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db 61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Qy 121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db 121 TTCAAGTCTAAGGACGTTCCCACTTCGCCCCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Qy 181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
Db 181 GTCAACCAGCCCGCCGTCGGTGCGGCAATGCGGCTGTTGAGACGGAATACTGCTTTTCAT 240
Qy 241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
Db 241 CGTGAAGATGGCACGGCAATTCCCGATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
Qy 301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
Db 301 GAAGGTGATGTTCTGTTTTTATACGGTTCAAAAGAAAATAAACTTCAACAACTTAAAAGC 360
Qy 361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
Db 361 GAAATTCATAAACGTAATCCTGAGGCAAGCATTACCACATCGGAAAATGAAAATAAAAAA 420
Qy 421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
Db 421 TATAATTATCGGTTTGTCAGTGCCGGTTATGTGTTTACTAAAAACGGAAAAGATGAAATT 480
Qy 481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
Db 481 GAGAAAACATCGGATGAAAAGCAGTTTTCTAATCGTTTAGGCTATGACGGTTTTGTATAT 540
Qy 541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
Db 541 TATCTCGGAGAACATCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
Qy 601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
Db 601 AACTGGCAATATATGACCGATGCCATACGTCATCGGAGAGGTAA---GGGGGTTTCCAGT 657
Qy 661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
Db 658 GTGGATTTGGGTTATACCACATATTATGGTAATGAAATTGGGGCAGCTTCTTATGAGGCT 717
Qy 721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
Db 718 AGGGATGCCGATGGCCGGGAAAAACATCCTGCCGAATATACGGTTAATTTCGACAAAAAA 777
Qy 781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
Db 778 AACCTGGAAGGTAAGTTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAT 837
Qy 841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
Db 838 CCACTGACCATTTACAACATTACCGCAACATTGGACGGCAACCGCTTTACCGGCAGTGCC 897
Qy 901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
Db 898 AAAGTTAGCACCGAGGTGAAGACGCAACACGCTGATAAAGAATATTTGTTTTTCCATACC 957
Qy 961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
Db 958 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAACGGAGAAGAGCTTGCCGGG 1017
Qy 1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
Db 1018 CGGTTTATCAGTAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1077
Qy 1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
Db 1078 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1137
Qy 1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
Db 1138 GATTCTCTAAAAATTTCCGTTGACGAGGCGACTGATGACCATGCCCGTAAGTTTGCCATT 1197
Qy 1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
Db 1198 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1257
Qy 1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
Db 1258 TTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCATCCGTGCT 1317
Qy 1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
Db 1318 TGTTGCGATTTTCTGACCTATGTGAAACTCGGACGGATAAAAACCGACCGCCCCGCCGTC 1377
Qy 1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
Db 1378 AAACCGAAGGCGCAGGATGAAGAGGATTCGGACATTGATAATGGCGAAGAAAGCGAAGAC 1437
Qy 1441 GAAA---GCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGA---------- 1481
Db 1438 GAAATTTCCGAAGATGATAACGGCGAAGATGAAGTCACCGAAGAAGAGGAAGCTGAAGAA 1497
Qy 1482 --CGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAA------GAA 1533
Db 1498 ACCGAAGAAGAAACTGATGAAGACGAAGAGGAAGAACCCGAAGAAACTGAAGAAACTGAA 1557
Qy 1534 GAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGC 1593
Db 1558 GAAACTGAAGAAACTGAAGAAACTGAAGAAACTGAAGAAAAATCGCCGACAGAAGAAGGC 1617
Qy 1594 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1653
Db 1618 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1677
Qy 1654 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCAAAAAACGGAACGGCGCAT 1713
Db 1678 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCGCAAATTGGAAAAGCACGC 1737
Qy 1714 TATACCGGCACTTGGGAAGCGCGTATCGGCGTATCGGATA---------- 1753
Db 1738 TATACCGGCACTTGGGAAGCGCGTATCGGCGTGCCGGATAAGAAAGGCGAACAGCTAGAT 1797
Qy 1754 -GTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAAGCAGAATTTACC 1812
Db 1798 GGCACTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGCGGCAAAAGCAGAATTTGAC 1857
Qy 1813 GTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAAATGATACAACCCCC 1872
Db 1858 GTTGATTTTGGTGCGAAGTCGCTTTCAGGTAAGTTGACAGAAAAAAAATGATACACACCCC 1917
Qy 1873 GCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCTTTGGCGCATACT 1932
Db 1918 GCTTTTTATATTGAAAAAGGTGTGATTGATGGCAACGGTTTCCACGCTTTGGCGCGTACT 1977
Qy 1933 CGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAAGCC 1992
Db 1978 CGTGAAAATGGTGTTGATTTGTCTGGGCAAGGTTCGACTAATCCCCAAAGTTTTAAAGCC 2037
Qy 1993 GACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGTAAT 2052
Db 2038 AGTAATCTTCTCGTAGAAGGAGGATTTTATGGTCCGCAGGCGGCAGAGTTGGGTGGTAAT 2097
Qy 2053 ATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGACAAGGAG 2112
Db 2098 ATTATCGACAGTGACCGGAAAATCGGCGTGGTATTCGGTGCGAAGAAAGATATGCAGGAG 2157
Qy 2113 GCAACACGATGA 2124
Db 2158 GTGGAAAAATGA 2169

```
RESULT 7
NMA6Z2491/c
LOCUS       NMA6Z2491             326301 bp    DNA     linear   BCT 02-SEP-2002
DEFINITION  Neisseria meningitidis serogroup A strain Z2491 complete genome;
            segment 6/7.
ACCESSION   AL162757 AL157959
VERSION     AL162757.2  GI:7380371
KEYWORDS    .
SOURCE      Neisseria meningitidis Z2491
  ORGANISM  Neisseria meningitidis Z2491
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE   1  (bases 1 to 326301)
  AUTHORS   Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
            Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
            Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
            Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
            Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
  TITLE     Complete DNA sequence of a serogroup A strain of Neisseria
            meningitidis Z2491
  JOURNAL   Nature 404 (6777), 502-506 (2000)
  MEDLINE   20222556
   PUBMED   10761919
REFERENCE   2  (bases 1 to 326301)
  AUTHORS   Parkhill,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
            sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of N. meningitidis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES             Location/Qualifiers
     source          1. .326301
                     /organism="Neisseria meningitidis Z2491"
                     /mol_type="genomic DNA"
                     /strain="Z2491"
                     /db_xref="taxon:122587"
                     /note="serogroup: A"
     repeat_unit     196. .311
                     /note="REP 2; hmmfs hit to HMM REP 2 (1 - 109), score:
                     87.43"
     repeat_unit     434. .453
                     /note=">= 90% match to ATTCCCNNNNNNNNGGGAAT"
                     /label=dRS3
     repeat_unit     498. .560
                     /note="Correia element; hmmfs hit to HMM Correia (1 - 62),
                     score: 78.71"
                     /label=Correia
     repeat_unit     562. .604
                     /note="Correia element; hmmfs hit to HMM Correia (114 -
                     156), score: 51.92"
                     /label=Correia
     misc_feature    complement(667. .676)
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     gene            complement(668. .1813)
                     /gene="dapE"
     CDS             complement(668. .1813)
                     /gene="dapE"
                     /EC_number="3.5.1.18"
                     /note="NMA1730, dapE, succinyl-diaminopimelate
                     desuccinylase, len: 381aa; similar to many eg. SW:P24176
                     (DAPE_ECOLI) dapE, succinyl-diaminopimelate desuccinylase
                     from Escherichia coli (375 aa) fasta scores; E(): 0, 56.9%
                     identity in 371 aa overlap. Contains Pfam match to entry
                     PF01546 Peptidase_M20."
                     /codon_start=1
                     /transl_table=11
                     /product="succinyl-diaminopimelate desuccinylase"
                     /protein_id="CAB84958.1"
                     /db_xref="GI:7380372"
                     /db_xref="GOA:Q9JTL0"
                     /db_xref="SPTREMBL:Q9JTL0"
                     /translation="MTETQSLELAKALISRPSVTPDDRDCQKLLVERLYKIGFAAEEL
                     HFGDTKNIWLRRGTKVPVVCFAGHTDVVPTGPVEKWDSPPFEPTERDGRLYGRGAADM
                     KTSIACFVTACERFVAEHPDHQGSIALLITSDEEGDALDGTTKVVDVLKARGELIDYC
                     IVGEPTAVDKLGDMIKNGRRGSLSGNLTVKGKQGHIAYPHLAINPVHTFAPALLELTQ
                     EVWDEGNKYFPPTSFQISNINGGTGATNVIPGELNVKFNFRFSTESTEAGLKQRVHAI
                     LDKHGVQYDLQWSCSGQPFLTQAGKLTDVARAAIAETCGIEAELSTTGGTSDGRFIKA
                     IAKELIELGPSNATIHQINENVRLDDIPKLSAVYEGILARLLAGNAV"
misc_feature         complement(836. .1786)
                     /gene="dapE"
                     /note="Pfam match to entry PF01546 Peptidase_M20, , score
                     252.50, E-value 5.7e-72"
misc_feature         complement(1863. .1872)
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
gene                 complement(1927. .2547)
                     /gene="NMA1731"
CDS                  complement(1927. .2547)
                     /gene="NMA1731"
                     /note="NMA1731, conserved hypothetical protein, len:
                     206aa; similar to SW:P44013 (Y552_HAEIN) hypothetical
                     protein from Haemophilus influenzae (207 aa) fasta scores;
                     E(): 0, 59.6% identity in 203 aa overlap."
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="CAB84959.1"
                     /db_xref="GI:7380373"
                     /db_xref="SPTREMBL:Q9JQZ3"
                     /translation="MLTPKSCDLFNIPPFQFSQLKKYQPESIPQIKADYKENWQIWQQ
                     LIQQVAAELGAPFAPPHIERWCNGWQVRAHFFAYFKYEQYKNSAAILSILLNRRRLSV
                     SLDWHCYKADVSPIALPDYNRWLDNFDTEKYASFDMWHGAESEYDDYRTVAQQNESDR
                     KLQNDEDFFCIGKHIERDDLGRQDVAKWIAETVEDLLPLYEACHGK"
RBS                  complement(2551. .2558)
misc_feature         complement(2570. .2579)
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
gene                 complement(2581. .3090)
                     /gene="NMA1732"
CDS                  complement(2581. .3090)
                     /gene="NMA1732"
                     /note="NMA1732, conserved hypothetical protein, len:
                     169aa; similar to SW:P43971 (Y243_HAEIN) hypothetical
                     protein from Haemophilus influenzae (172 aa) fasta scores;
                     E(): 0, 49.7% identity in 169 aa overlap."
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="CAB84960.1"
                     /db_xref="GI:7380374"
                     /db_xref="SPTREMBL:Q9JTK9"
                     /translation="MNPPETQSVTFAEPIEMLYACHGKVRRFCGQVAMLSDYIAENGC
                     NQLVLQTIRQITRYFNVAAPLHHEDEEENFFPLLLQYAPQAQESVDELLRQHVSLHGN
                     WDAVAAEFAKLEADNAYIPDAEAFKRFVAGYDVHLAIEEPLFDMGKMFIPKEKLTEIG
                     EIMAARRRR"
gene                 complement(3185. .3736)
                     /gene="NMA1733"
CDS                  complement(3185. .3736)
                     /gene="NMA1733"
                     /note="NMA1733, H.8 outer membrane protein, len: 183aa;
                     strongly similar to SW:P07212 (H8_NEIME) H.8 outer
                     membrane protein from Neisseria meningitidis (183 aa)
                     fasta scores; E(): 0, 96.7% identity in 183 aa overlap.
                     Contains Pfam match to entry PF00127 copper-bind, Copper
                     binding proteins, plastocyanin/azurin family and Prosite
                     match to PS00196 Type-1 copper (blue) proteins signature."
                     /codon_start=1
                     /transl_table=11
                     /product="H.8 outer membrane protein"
                     /protein_id="CAB84961.1"
```

/db_xref="GI:7380375"
/db_xref="GOA:P57025"
/db_xref="SWISS-PROT:P57025"
/translation="MKAYLALISAAVIGLAACSQEPAAPAAEATPAAEAPASEAPAAE
AAPADAAEAPAAGNCAATVESNDNMQFNTKDIQVSKACKEFTITLKHTGTQPKTSMGH
NIVIGKTEDMDGIFKDGVGAADTDYVKPDDARVVAHTKLIGGGEESSLTLDPAKLADG
EYKFACTFPGHGALMNGKVTLVD"
misc_feature complement(3191. .3568)
/gene="NMA1733"
/note="Pfam match to entry PF00127 copper-bind, Copper binding proteins, plastocyanin/azurin family, score 218.60, E-value 9.1e-62"
misc_feature complement(3212. .3259)
/gene="NMA1733"
/note="PS00196 Type-1 copper (blue) proteins signature."
gene complement(4017. .4439)
/gene="NMA1734"
CDS complement(4017. .4439)
/gene="NMA1734"
/note="NMA1734, unknown, len: 140aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB84962.1"
/db_xref="GI:7380376"
/db_xref="SPTREMBL:Q9JTK8"
/translation="MNLEQLGRRDALLSGLLKQAGQWRRLDAAVKKLLPANLHPHFQT
ACIEDGRLVLLAANNMAASRLKMIAPSVLPQLAGLDASIRSVSVRLVPKPEKPPKTNT
LHLSKAALESFGSAAVKLEKRHPELAEALANLVRRHGA"
misc_feature 4306. .4315
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
misc_feature complement(4486. .4495)
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
gene 4570. .7320
/gene="secA"
CDS 4570. .7320
/gene="secA"
/note="NMA1735, secA, preprotein translocase SecA subunit, len: 916aa; similar to many eg. SW:P10408 (SECA_ECOLI) preprotein translocase SecA subunit from Escherichia coli (901 aa) fasta scores; E(): 0, 58.0% identity in 920 aa overlap. Contains Pfam match to entry PF01043 SecA_protein, SecA protein, amino terminal region and Prosite match to PS01312 Protein secA signatures."
/codon_start=1
/transl_table=11
/product="preprotein translocase SecA subunit"
/protein_id="CAB84963.1"
/db_xref="GI:7380377"
/db_xref="GOA:Q9JTK7"
/db_xref="SPTREMBL:Q9JTK7"
/translation="MLTNIAKKIFGSRNDRLLKQYRKSVARINALEEQMQALSDADLQ
AKTAEFKQRLADGQTLDGILPEAFAVCREASRRTLGMRHFDVQLIGGMVLHDGKIAEM
RTGEGKTLVATLAVYLNALAGKGVHVVTVNDYLASRDAGIMEPLYNFLGLTVGVIISD
MQPFDRQNAYAADITYGTNNEFGFDYLRDNMVTDQYDKVQRELNFAVVDEVDSILIDE

Query Match 71.4%; Score 1516.8; DB 1; Length 326301;
Best Local Similarity 83.3%; Pred. No. 3.2e-307;
Matches 1792; Conservative 0; Mismatches 307; Indels 51; Gaps 4;

```
Qy        1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db    18981 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGTATC 18922

Qy       61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db    18921 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 18862

Qy      121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db    18861 TTCAAGTCTAAGGACGTTCCCACTTCGCCCCCTGCCGGGCCTTCGGTAGAAACCACGCCG 18802

Qy      181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
Db    18801 GTCAACCGGCCCGCCGTCGGTGCGGCAATGCGGCTGTTGAGGCGGAATATTGCTTTTCAT 18742

Qy      241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
Db    18741 CGTGAAGATGGCACGGCAATTCCCGATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 18682

Qy      301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
Db    18681 GAAGGTGATGTTCTGTTTTTATACGGTTCCAAAAAAGATAAACTTCAGCAGCTTAAGGAT 18622

Qy      361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
Db    18621 AAAATTCATCAACGCAATCCTAATGTAGAAATTAGGACATCAGAAAATGAAAATAAAAAA 18562

Qy      421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
Db    18561 TATGGTTATGAATTTGTAGATGCAGGTTATGTATATACTACAAAGGGAAAAGATGAAATT 18502

Qy      481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
Db    18501 GAGTGGACTTCAAATCACAAGCAGTTTACCTACCGGTTTGGTTATGACGGTTTTGTATAT 18442

Qy      541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
Db    18441 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 18382

Qy      601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
Db    18381 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 18322

Qy      661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
Db    18321 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 18262

Qy      721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
Db    18261 ACTGCCGACGACCGAGAGGGAAAACATCCTGCCGAATATACGGTAGATTTCGGTAAGAAA 18202

Qy      781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
Db    18201 ACTTTGACGGGTAAATTAATTAAAAATCAGTATGTGCAAAAGAAAACCGATGAAAAGAAA 18142

Qy      841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
Db    18141 CCGCTGACCATTTACGACATTACTGCAACATTGGACGGCAACCGCTTTACCGGCAGTGCC 18082

Qy      901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
Db    18081 AAAGTTAACACCGAGGTGAAGACGAAACACGCTGATAAAGAGCATTTGTTTTTCCATACC 18022

Qy      961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
Db    18021 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAACGGAGAAGAGCTTGCCGGA 17962

Qy     1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
Db    17961 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTATTCGCAGGCAAACAAAAAACAGAC 17902

Qy     1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
Db    17901 GCATCAAACGCATCAGATACAAATCCTGCTATGCCGTCTGAAAAACACACCAAAATCTTG 17842

Qy     1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
Db    17841 GATTCTCTGAAAATTTCCGTTGACGAGGCGACTGATAGCAATGCCCGTAAGTTTGCCATT 17782

Qy     1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
Db    17781 TCCCCTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATCTCC 17722
```

```
Qy       1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
Db      17721 TTGCTCAAAGATACTCAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCAGTGCT 17662

Qy       1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
Db      17661 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGACCGCCCGGCAAGT 17602

Qy       1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGC--------------- 1425
Db      17601 AAACCAAAGGCGGAAGATGAAAATTCCGAAGATGAAATTGGTGAAAGCGAGGAAAATGAA 17542

Qy       1426 ---------GTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACC 1476
Db      17541 GAGGATTTGGTCGCTGAAGAAGAAAACACGGAAGACGAAGTCGTAGAAGATGAAGACAGA 17482

Qy       1477 GGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAA 1536
Db      17481 GAAGAAGACGAAGTTTCCGAAGATGGTAACAGTGAAGACGAAGAAGAAATCGCCGAAGAA 17422

Qy       1537 CCTGAAGAAGAAGCTGAAGAG---------GAAGAACCCGAAGAAGAATTGCCGGCAGAA 1587
Db      17421 GATGATGATGAAGCCGAAGAGGAGGAAGTTGAAGAACCCGAAGAAGAATCGCCGGAAGAA 17362

Qy       1588 GAAGGCAACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGAC 1647
Db      17361 GGCGGCGGTGGCGGTTCAGACGGCATCCCGCCCGCTTCGGAAGCCCCTAAAGGCAGGGAC 17302

Qy       1648 ATCGACCTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCAAAAAACGGAACG 1707
Db      17301 ATCGACCTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCGCAAACTGGAAAA 17242

Qy       1708 GCGCATTATACCGGCACTTGGGAAGCGCGTATCGGCGTATCGGATAGTGGTACGTCCATT 1767
Db      17241 GCACGCTATACCGGCACTTGGGAAGCGCGTATCAGC---------------AAACCCATT 17197

Qy       1768 CAAAAGGATAGCTATGCGAATCAA---GGGGCAAAAGCAGAATTTACCGTTGATTTCGAA 1824
Db      17196 CAATGGGATAATAAGGCGGATAAAAAAGCGGCAAAAGCAGAATTTGACGTTGATTTCGGC 17137

Qy       1825 GCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAATGATACAACCCCCGCTTTTTATATT 1884
Db      17136 GAGAAATCGATTTCCGGAACGCTGACGGAGAAAAACGGTGTAGAAGCTGCTTTCTATATT 17077

Qy       1885 GAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCTTTGGCGCATACTCGGGAGAACGGT 1944
Db      17076 GAAAAAGGTGTGATTGATGGCAACGGTTTCCACGCGACAGCACGCACTCGGGATAACGGC 17017

Qy       1945 ATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAAGCCGACAATCTTCTT 2004
Db      17016 ATCAATCTTTCGGGAAATGGTTCGACCAACCCCAAAACCTTCCAAGCTAGTGATCTTCGT 16957

Qy       2005 GTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGTAATATTATCGACAGC 2064
Db      16956 GTAGAAGGAGGATTTTACGGCCCGCAGGCGGAGGAATTGGGCGGTACTATTTTCAATAAG 16897

Qy       2065 GACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGACAAGGAGGC 2114
Db      16896 GATGGGAAATCTCTTGATATAACTGAAGATATTGACAATGAAATTGAAGC 16847
```

RESULT 8
AF123380
LOCUS AF123380 2226 bp DNA linear BCT 24-MAY-1999
DEFINITION Neisseria meningitidis strain H44/76 lactoferrin-binding protein precursor (lbpB) gene, complete cds.
ACCESSION AF123380
VERSION AF123380.1 GI:4884686
KEYWORDS .
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.

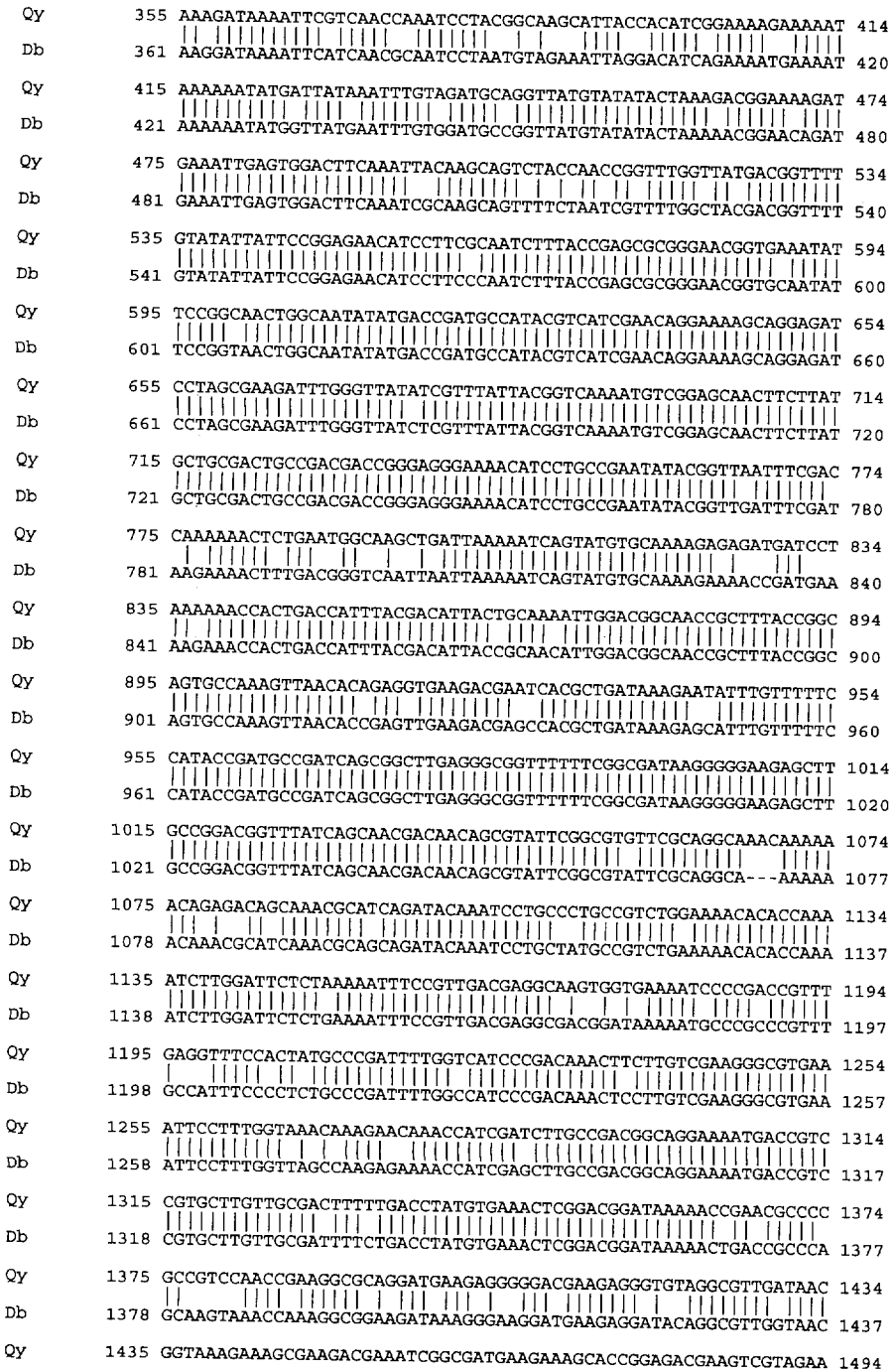
REFERENCE 1 (bases 1 to 2226)
AUTHORS Pettersson,A., van der Biezen,J., Joosten,V., Hendriksen,J. and Tommassen,J.
TITLE Sequence variability of the meningococcal lactoferrin-binding protein LbpB
JOURNAL Gene 231 (1-2), 105-110 (1999)
MEDLINE 99250255
PUBMED 10231574
REFERENCE 2 (bases 1 to 2226)
AUTHORS Pettersson,A., van der Biezen,J., Joosten,V., Hendriksen,J. and Tommassen,J.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1999) Department of Molecular Cell Biology, Utrecht University, Padualaan 8, Utrecht 3584 CH, The Netherlands

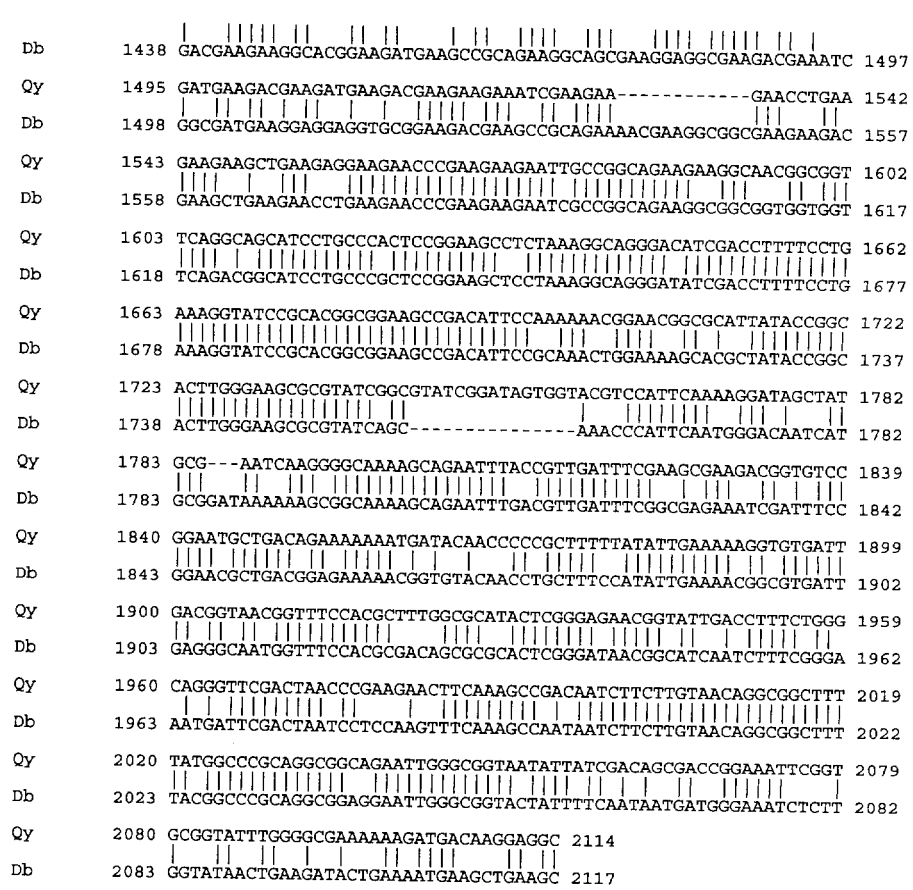
FEATURES Location/Qualifiers
source 1. .2226
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="H44/76"
/db_xref="taxon:487"
gene 1. .2226
/gene="lbpB"
CDS 1. .2226
/gene="lbpB"
/codon_start=1
/transl_table=11
/product="lactoferrin-binding protein precursor"
/protein_id="AAD31768.1"
/db_xref="GI:4884687"
/translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
DVPTPPPAKPSIETTPVPSTGPAVGAAMRLLRRIFATSDKVGNDFPNSKQAEEKLSFK
EGDVLFLYGSKKDKLQWLKDKIHQRNPNVEIRTSENENKKYGYEFVDAGYVYTKNGTD
EIEWTSNRKQFSNRFGYDGFVYYSGEHPSQSLPSAGTVQYSGNWQYMTDAIRHRTGKA
GDPSEDLGYLVYYGQNVGATSYAATADDREGKHPAEYTVDFDKKTLTGQLIKNQYVQK
KTDEKKPLTIYDITATLDGNRFTGSAKVNTELKTSHADKEHLFFHTDADQRLEGGFFG
DKGEELAGRFISNDNSVFGVFAGKKTNASNAADTNPAMPSEKHTKILDSLKISVDEAT
DKNARPFAISPLPDFGHPDKLLVEGREIPLVSQEKTIELADGRKMTVRACCDFLTYVK
LGRIKTDRPASKPKAEDKGKDEEDTGVGNDEEGTEDEAAEGSEGGEDEIGDEGGGAED
EAAENEGGEEDEAEEPEEPEEESPAEGGGGGSDGILPAPEAPKGRDIDLFLKGIRTAE
ADIPQTGKARYTGTWEARISKPIQWDNHADKKAAKAEFDVDFGEKSISGTLTEKNGVQ
PAFHIENGVIEGNGFHATARTRDNGINLSGNDSTNPPSFKANNLLVTGGFYGPQAEEL
GGTIFNNDGKSLGITEDTENEAEAEVENEAGVGEQLKPEAKPQFGVVFGAKKDNKEVE
K"

ORIGIN

Query Match 70.6%; Score 1499.8; DB 1; Length 2226;
Best Local Similarity 83.3%; Pred. No. 7.6e-304;
Matches 1779; Conservative 0; Mismatches 317; Indels 39; Gaps 5;

```
Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGTATT 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db        121 TTCAAGTCTAAGGACGTTCCCACTCCGCCCCCTGCCAAACCTTCTATAGAAACCACGCCG 180

Qy        181 GT------CAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCA 234
Db        181 GTGCCGTCAACCGGGCCTGCCGTCGGTGCGGCAATGCGGCTGTTGAGGCGGATTTTCGCA 240

Qy        235 ACTTCTGATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCG 294
Db        241 ACTTCTGATAAGGTTGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCG 300

Qy        295 TTTAAAGAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAAGATCAACGTCAGCAGCTT 354
Db        301 TTTAAAGAAGGTGATGTTCTGTTTTTATACGGTTCAAAAAAAGATAAACTTCAGTGGCTT 360
```





RESULT 9
A98972
LOCUS A98972 2226 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 5 from Patent WO9909176.
ACCESSION A98972
VERSION A98972.1 GI:6781932
KEYWORDS .
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 2226)
AUTHORS Pettersson-Fernholm,A.M. and Tommassen,J.P.
TITLE NEISSERIA LACTOFERRIN BINDING PROTEIN
JOURNAL Patent: WO 9909176-A 5 25-FEB-1999;
UNIV UTRECHT (NL); PETTERSSON FERNHOLM ANNIKA MAR (NL)
FEATURES Location/Qualifiers
source 1. .2226
/organism="Neisseria meningitidis"
/mol_type="unassigned DNA"
/strain="H44/76"
/db_xref="taxon:487"
CDS 1. .2226
/note="unnamed protein product"
/codon_start=1

```
                     /transl_table=11
                     /protein_id="CAB70583.1"
                     /db_xref="GI:6781933"
                     /db_xref="REMTREMBL:CAB70583"
                     /translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
                     DVPTPPPAKPSIETTPVPSTGPAVGAAMRLLRRIFATSDKVGNDFPNSKQAEEKLSFK
                     EGDVLFLYGSKKDKLQWLKDKIHQRNPNVEIRTSENENKKYGYEFVDAGYVYTKNGTD
                     EIEWTSNRKQFSNRFGYDGFVYYSGEHPSQSLPSAGTVQYSGNWQYMTDAIRHRTGKA
                     GDPSEDLGYLVYYGQNVGATSYAATADDREGKHPAEYTVDFDKKTLTGQLIKNQYVQK
                     KTDEKKPLTIYDITATLDGNRFTGSAKVNTELKTSHADKEHLFFHTDADQRLEGGFFG
                     DKGEELAGRFISNDNSVFGVFAGKKTNASNAADTNPAMPSEKHTKILDSLKISVDEAT
                     DKNARPFAISPLPDFGHPDKLLVEGREIPLVSQEKTIELADGRKMTVRACCDFLTYVK
                     LGRIKTDRPASKPKAEDKGKDEEDTGVGNDEEGTEDEAAEGSEGGEDEIGDEGGGAED
                     EAAENEGGEEDEAEEPEEPEEESPAEGGGGGSDGILPAPEAPKGRDIDLFLKGIRTAE
                     ADIPQTGKARYTGTWEARISKPIQWDNHADKKAAKAEFDVDFGEKSISGTLTEKNGVQ
                     PAFHIENGVIEGNGFHATARTRDNGINLSGNDSTNPPSFKANNLLVTGGFYGPQAEEL
                     GGTIFNNDGKSLGITEDTENEAEAEVENEAGVGEQLKPEAKPQFGVVFGAKKDNKEVE
                     K"
ORIGIN

  Query Match             70.6%;  Score 1499.8;  DB 6;  Length 2226;
  Best Local Similarity   83.3%;  Pred. No. 7.6e-304;
  Matches 1779;  Conservative    0;  Mismatches  317;  Indels   39;  Gaps    5;

Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGTATT 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
              |||||||||||||||||||||||| |||| ||||||     ||||  ||||||||||||||
Db        121 TTCAAGTCTAAGGACGTTCCCACTCCGCCCCCTGCCAAACCTTCTATAGAAACCACGCCG 180

Qy        181 GT------CAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCA 234
              ||      | | | || ||||| |||||||||||||||||||||| |||| | | |||
Db        181 GTGCCGTCAACCGGGCCTGCCGTCGGTGCGGCAATGCGGCTGTTGAGGCGGATTTTCGCA 240

Qy        235 ACTTCTGATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCG 294
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTTCTGATAAGGTTGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCG 300

Qy        295 TTTAAAGAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTT 354
              |||||||| |  ||| | |||||||||||||||| ||||||||| ||| ||||  ||||
Db        301 TTTAAAGAAGGTGATGTTCTGTTTTTATACGGTTCAAAAAAAGATAAACTTCAGTGGCTT 360

Qy        355 AAAGATAAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAAT 414
              || |||||||||| |||||  |||||||  | |   ||||  ||||| |||||  |||||
Db        361 AAGGATAAAATTCATCAACGCAATCCTAATGTAGAAATTAGGACATCAGAAAATGAAAAT 420

Qy        415 AAAAAATATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGAT 474
              |||||||||| |||| |||||| ||||| |||||||||||||||||||| |||||| ||||
Db        421 AAAAAATATGGTTATGAATTTGTGGATGCCGGTTATGTATATACTAAAAACGGAACAGAT 480

Qy        475 GAAATTGAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTT 534
              |||||||||||||||||||  |||||||  | | || || ||||| || |||||||||
Db        481 GAAATTGAGTGGACTTCAAATCGCAAGCAGTTTTCTAATCGTTTTGGCTACGACGGTTTT 540

Qy        535 GTATATTATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATAT 594
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        541 GTATATTATTCCGGAGAACATCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGCAATAT 600

Qy        595 TCCGGCAACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGAT 654
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCGGTAACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGAT 660

Qy        655 CCTAGCGAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTAT 714
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        661 CCTAGCGAAGATTTGGGTTATCTCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTAT 720

Qy        715 GCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGAC 774
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        721 GCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTGATTTCGAT 780

Qy        775 CAAAAAACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCT 834
              | |||||| ||| ||  |  | |||||||||||||||||||||||||| |   |||
Db        781 AAGAAAACTTTGACGGGTCAATTAATTAAAAATCAGTATGTGCAAAAGAAAACCGATGAA 840

Qy        835 AAAAAACCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGC 894
              || |||||||||||||||||||||||||| |||| ||||||||||||||||||||||||
Db        841 AAGAAACCACTGACCATTTACGACATTACCGCAACATTGGACGGCAACCGCTTTACCGGC 900

Qy        895 AGTGCCAAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTC 954
              ||||||||||||||||| ||| |||||||||  |||||||||||||  ||||||||||||
Db        901 AGTGCCAAAGTTAACACCGAGTTGAAGACGAGCCACGCTGATAAAGAGCATTTGTTTTTC 960

Qy        955 CATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTT 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTT 1020

Qy       1015 GCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAA 1074
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1021 GCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTATTCGCAGGCA---AAAAA 1077

Qy       1075 ACAGAGACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAA 1134
              ||| |  || |||||||| ||||||||||||||||||||||||| |||||||||||||||
Db       1078 ACAAACGCATCAAACGCAGCAGATACAAATCCTGCTATGCCGTCTGAAAAACACACCAAA 1137

Qy       1135 ATCTTGGATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTT 1194
              |||||||||||||||||||||||||||||||||||||||||| |  | | ||||| |||| ||||||
Db       1138 ATCTTGGATTCTCTGAAAATTTCCGTTGACGAGGCGACGGATAAAAATGCCCGCCCGTTT 1197

Qy       1195 GAGGTTTCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAA 1254
              |   ||||| || |||||||||||||| |||||||||||||||||||||||||||||||
Db       1198 GCCATTTCCCCTCTGCCCGATTTTGGCCATCCCGACAAACTCCTTGTCGAAGGGCGTGAA 1257

Qy       1255 ATTCCTTTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTC 1314
              |||||||||| | |||| |||||||||| ||||||||||||||||||||||||||||||
Db       1258 ATTCCTTTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCGTC 1317

Qy       1315 CGTGCTTGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCC 1374
              |||||||||||||| ||| ||||||||||||||||||||||||||||||| || |||||
Db       1318 CGTGCTTGTTGCGATTTTCTGACCTATGTGAAACTCGGACGGATAAAAACTGACCGCCCA 1377

Qy       1375 GCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAAC 1434
              ||     |||| |||||| | ||| ||| |  ||| ||||||| | |||||||| ||||
Db       1378 GCAAGTAAACCAAAGGCGGAAGATAAAGGGAAGGATGAAGAGGATACAGGCGTTGGTAAC 1437

Qy       1435 GGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAA 1494
              |  |||||||   || ||  | | |||| |||  |||| ||||| || |
Db       1438 GACGAAGAAGGCACGGAAGATGAAGCCGCAGAAGGCAGCGAAGGAGGCGAAGACGAAATC 1497

Qy       1495 GATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAA----------GAACCTGAA 1542
              |  || || | | ||  |  | ||||| |||  || ||||              |||   ||
Db       1498 GGCGATGAAGGAGGAGGTGCGGAAGACGAAGCCGCAGAAAACGAAGGCGGCGAAGAAGAC 1557

Qy       1543 GAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGT 1602
              ||||  |  |||  |||||||||||||||||| |||||||||  |||  || |||
Db       1558 GAAGCTGAAGAACCTGAAGAACCCGAAGAAGAATCGCCGGCAGAAGGCGGCGGTGGTGGT 1617

Qy       1603 TCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTG 1662
              |||| | |||||||||| |||||||||||  |||||||||||||||||||||||||||||
Db       1618 TCAGACGGCATCCTGCCCGCTCCGGAAGCTCCTAAAGGCAGGGATATCGACCTTTTCCTG 1677

Qy       1663 AAAGGTATCCGCACGGCGGAAGCCGACATTCCAAAAAACGGAACGGCGCATTATACCGGC 1722
              |||||||||||||||||||||||||||||||  |||  ||||  || |  |||||||||
Db       1678 AAAGGTATCCGCACGGCGGAAGCCGACATTCCGCAAACTGGAAAAGCACGCTATACCGGC 1737

Qy       1723 ACTTGGGAAGCGCGTATCGGCGTATCGGATAGTGGTACGTCCATTCAAAAGGATAGCTAT 1782
              |||||||||||||||||| ||              |  ||||||| ||| |  ||
Db       1738 ACTTGGGAAGCGCGTATCAGC--------------AAACCCATTCAATGGGACAATCAT 1782

Qy       1783 GCG---AATCAAGGGGCAAAAGCAGAATTTACCGTTGATTTCGAAGCGAAGACGGTGTCC 1839
              |||   ||  ||| ||||||||||||||||  ||||||||||  | |||  || | |||
```

```
Db       1783 GCGGATAAAAAAGCGGCAAAAGCAGAATTTGACGTTGATTTCGGCGAGAAATCGATTTCC 1842
Qy       1840 GGAATGCTGACAGAAAAAAATGATACAACCCCCGCTTTTTATATTGAAAAAGGTGTGATT 1899
Db       1843 GGAACGCTGACGGAGAAAAACGGTGTACAACCTGCTTTCCATATTGAAAACGGCGTGATT 1902
Qy       1900 GACGGTAACGGTTTCCACGCTTTGGCGCATACTCGGGAGAACGGTATTGACCTTTCTGGG 1959
Db       1903 GAGGGCAATGGTTTCCACGCGACAGCGCGCACTCGGGATAACGGCATCAATCTTTCGGGA 1962
Qy       1960 CAGGGTTCGACTAACCCGAAGAACTTCAAAGCCGACAATCTTCTTGTAACAGGCGGCTTT 2019
Db       1963 AATGATTCGACTAATCCTCCAAGTTTCAAAGCCAATAATCTTCTTGTAACAGGCGGCTTT 2022
Qy       2020 TATGGCCCGCAGGCGGCAGAATTGGGCGGTAATATTATCGACAGCGACCGGAAATTCGGT 2079
Db       2023 TACGGCCCGCAGGCGGAGGAATTGGGCGGTACTATTTTCAATAATGATGGGAAATCTCTT 2082
Qy       2080 GCGGTATTTGGGGCGAAAAAAGATGACAAGGAGGC 2114
Db       2083 GGTATAACTGAAGATACTGAAAATGAAGCTGAAGC 2117
```

RESULT 10
BD074757

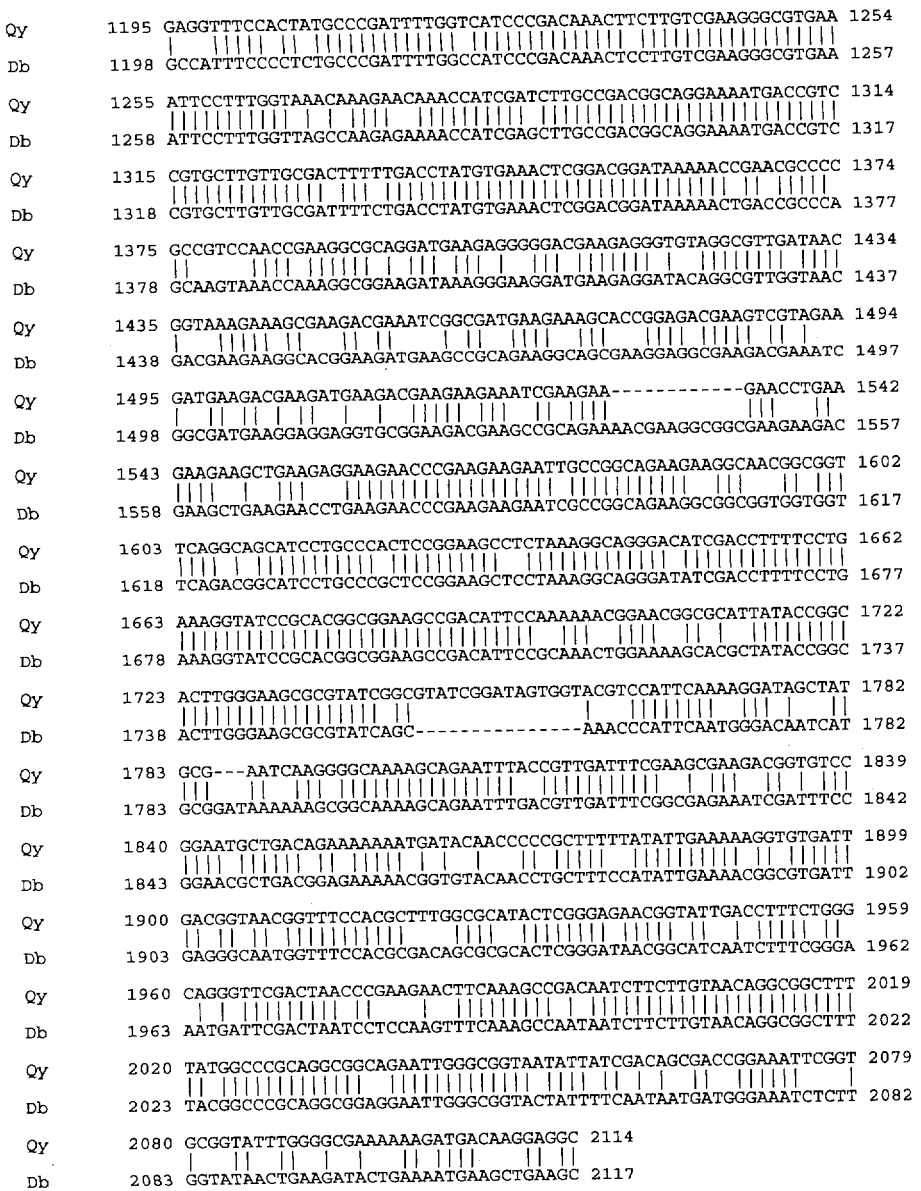
LOCUS	BD074757 2226 bp DNA linear PAT 27-AUG-2002
DEFINITION	Neisseria lactoferrin-binding protein.
ACCESSION	BD074757
VERSION	BD074757.1 GI:22620360
KEYWORDS	JP 2001514894-A/3.
SOURCE	unidentified
ORGANISM	unidentified
	unclassified.
REFERENCE	1 (bases 1 to 2226)
AUTHORS	Felnholm,A.M.P. and Thomasen,J.P.M.
TITLE	Neisseria lactoferrin-binding protein
JOURNAL	Patent: JP 2001514894-A 3 18-SEP-2001;
	UNIVERSITY OF UTRECHT,TECHNOLOGY FOUNDATION
COMMENT	OS Unidentified
	PN JP 2001514894-A/3
	PD 18-SEP-2001
	PF 10-AUG-1998 JP 2000509840
	PR 15-AUG-1997 GB 9717423.9,05-FEB-1998 GB 9802544.8 PI
	ANICA MARGARETA PETERSON FELNHOLM,JOHANES PETRUS MARIA PI
	THOMASEN
	PC C12N15/09,A61K39/095,A61K39/395,A61K48/00,A61P31/12,C07K14/22,
	PC C07K16/12,
	PC C12N1/21,C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/68, PC
	C12N15/00
	CC Strandedness: Double;
	CC Topology: Linear;
	CC Neisseria lactoferrin-binding protein
	FH Key Location/Qualifiers
	FT source 1. .2226
	FT /organism='Unidentified'.

FEATURES Location/Qualifiers
source 1. .2226
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 70.6%; Score 1499.8; DB 6; Length 2226;
Best Local Similarity 83.3%; Pred. No. 7.6e-304;
Matches 1779; Conservative 0; Mismatches 317; Indels 39; Gaps 5;

```
Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGTATT 60
Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db        121 TTCAAGTCTAAGGACGTTCCCACTCCGCCCCCTGCCAAACCTTCTATAGAAACCACGCCG 180
Qy        181 GT------CAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCA 234
Db        181 GTGCCGTCAACCGGGCCTGCCGTCGGTGCGGCAATGCGGCTGTTGAGGCGGATTTTCGCA 240
Qy        235 ACTTCTGATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCG 294
Db        241 ACTTCTGATAAGGTTGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCG 300
Qy        295 TTTAAAGAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTT 354
Db        301 TTTAAAGAAGGTGATGTTCTGTTTTTATACGGTTCAAAAAAAGATAAACTTCAGTGGCTT 360
Qy        355 AAAGATAAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAAT 414
Db        361 AAGGATAAAATTCATCAACGCAATCCTAATGTAGAAATTAGGACATCAGAAAATGAAAAT 420
Qy        415 AAAAAATATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGAT 474
Db        421 AAAAAATATGGTTATGAATTTGTGGATGCCGGTTATGTATATACTAAAAACGGAACAGAT 480
Qy        475 GAAATTGAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTT 534
Db        481 GAAATTGAGTGGACTTCAAATCGCAAGCAGTTTTCTAATCGTTTTGGCTACGACGGTTTT 540
Qy        535 GTATATTATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATAT 594
Db        541 GTATATTATTCCGGAGAACATCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGCAATAT 600
Qy        595 TCCGGCAACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGAT 654
Db        601 TCCGGTAACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGAT 660
Qy        655 CCTAGCGAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTAT 714
Db        661 CCTAGCGAAGATTTGGGTTATCTCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTAT 720
Qy        715 GCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGAC 774
Db        721 GCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTGATTTCGAT 780
Qy        775 CAAAAAACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCT 834
Db        781 AAGAAAACTTTGACGGGTCAATTAATTAAAAATCAGTATGTGCAAAAGAAAACCGATGAA 840
Qy        835 AAAAAACCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGC 894
Db        841 AAGAAACCACTGACCATTTACGACATTACCGCAACATTGGACGGCAACCGCTTTACCGGC 900
Qy        895 AGTGCCAAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTC 954
Db        901 AGTGCCAAAGTTAACACCGAGTTGAAGACGAGCCACGCTGATAAAGAGCATTTGTTTTTC 960
Qy        955 CATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTT 1014
Db        961 CATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTT 1020
Qy       1015 GCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAA 1074
Db       1021 GCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTATTCGCAGGCA---AAAAA 1077
Qy       1075 ACAGAGACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAA 1134
Db       1078 ACAAACGCATCAAACGCAGCAGATACAAATCCTGCTATGCCGTCTGAAAAACACACCAAA 1137
Qy       1135 ATCTTGGATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTT 1194
Db       1138 ATCTTGGATTCTCTGAAAATTTCCGTTGACGAGGCGACGGATAAAAATGCCCGCCCGTTT 1197
```



RESULT 11
AF072890
LOCUS AF072890 2519 bp DNA linear BCT 22-JAN-1999
DEFINITION Neisseria gonorrhoeae lactoferrin binding protein B precursor (lbpB) gene, complete cds.
ACCESSION AF072890
VERSION AF072890.1 GI:4106392
KEYWORDS .

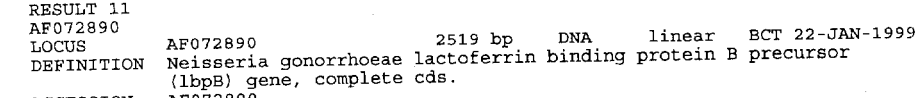
SOURCE Neisseria gonorrhoeae
ORGANISM Neisseria gonorrhoeae
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 2519)
AUTHORS Biswas,G.D., Anderson,J.E., Chen,C.J., Cornelissen,C.N. and Sparling,P.F.
TITLE Identification and functional characterization of the Neisseria gonorrhoeae lbpB gene product
JOURNAL Infect. Immun. 67 (1), 455-459 (1999)
MEDLINE 99081783
PUBMED 9864256
REFERENCE 2 (bases 1 to 2519)
AUTHORS Biswas,G.D., Anderson,J.E., Cornellisen,C.N. and Sparling,P.F.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1998) Medicine/ID, Univ. of N. Carolina, 521 Burnett Womack/CB 7030, Chapel Hill, N.C. 27599, USA

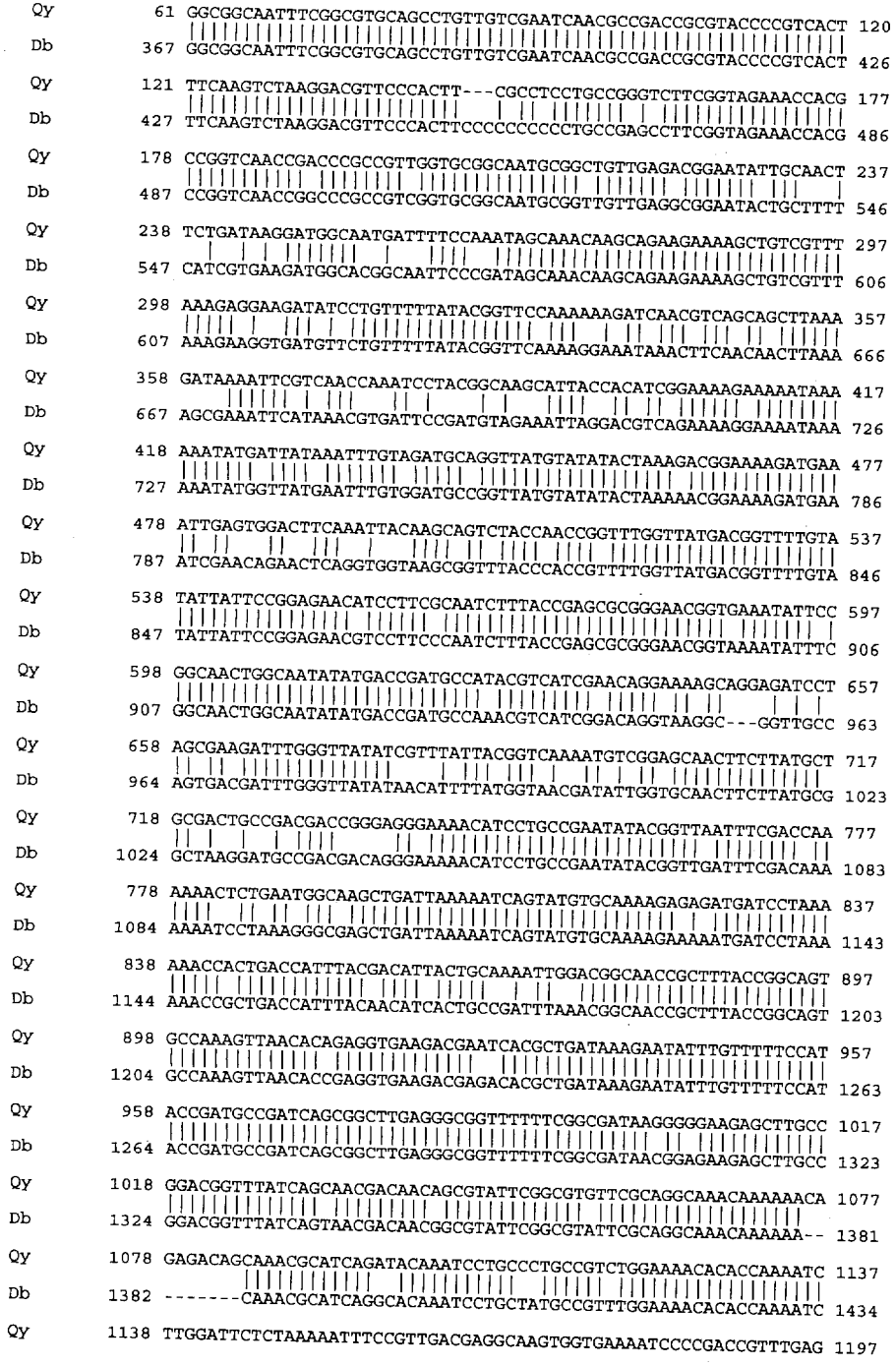
FEATURES Location/Qualifiers

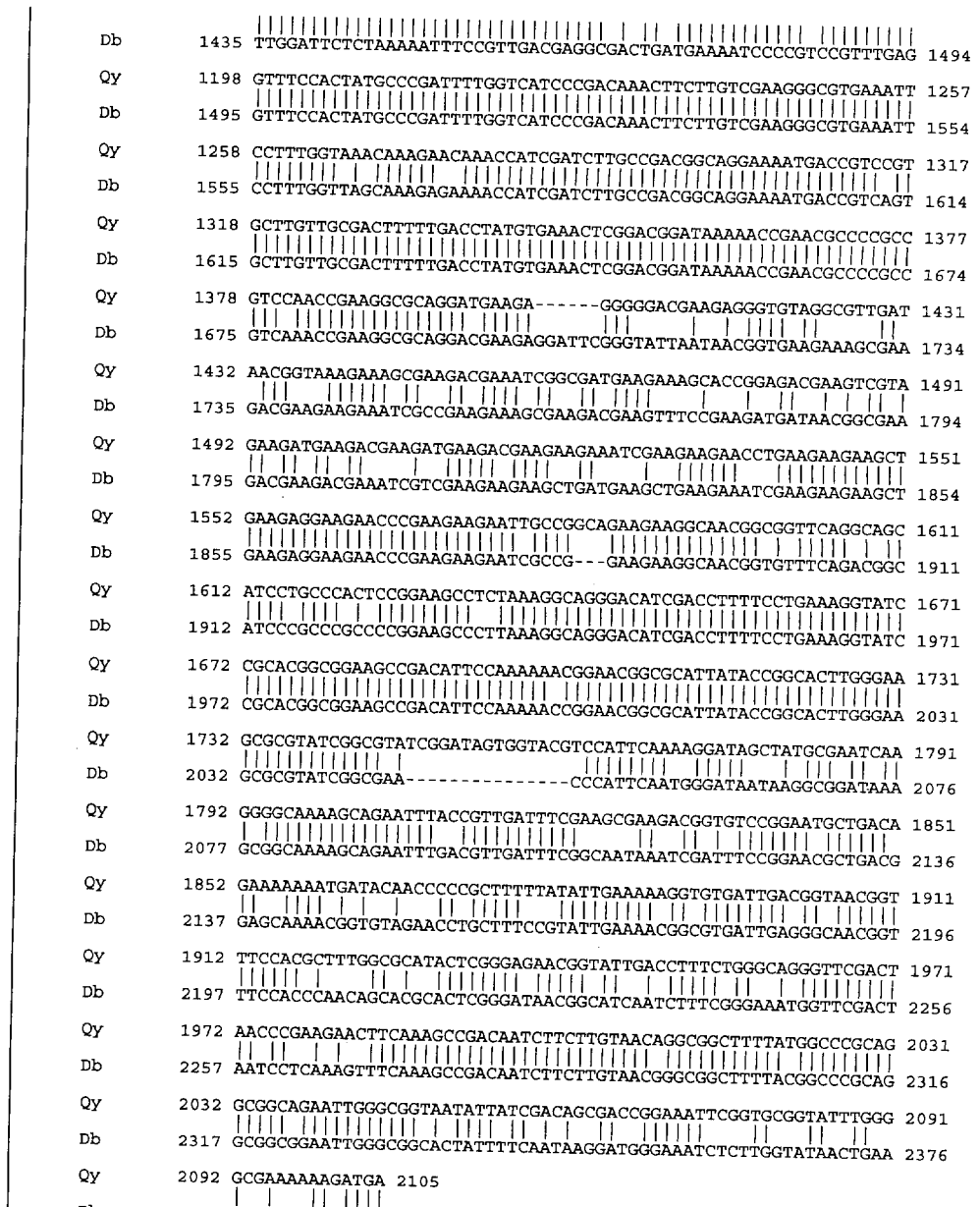
source	1. .2519
	/organism="Neisseria gonorrhoeae"
	/mol_type="genomic DNA"
	/strain="FA19"
	/db_xref="taxon:485"
terminator	184. .209
	/note="putative"
gene	276. .2496
	/gene="lbpB"
protein_bind	276. .293
	/gene="lbpB"
	/note="putative"
	/bound_moiety="FUR"
RBS	301. .304
	/gene="lbpB"
CDS	310. .2496
	/gene="lbpB"
	/note="LbpB"
	/codon_start=1
	/transl_table=11
	/product="lactoferrin binding protein B precursor"
	/protein_id="AAD08809.1"
	/db_xref="GI:4106393"
	/translation="MRKLNYYGIALLPLMLASCGGNFGVQPVVESTPTAYPVTFKSKD VPTSPPPAEPSVETTPVNRPAVGAAMRLLRRNTAFHREDGTAIPDSKQAEEKLSFKEG DVLFLYGSKGNKLQQLKSEIHKRDSDVEIRTSEKENKKYGYEFVDAGYVYTKNGKDEI EQNSGGKRFTHRFGYDGFVYYSGERPSQSLPSAGTVKYFGNWQYMTDAKRHRTGKAVA SDDLGYITFYGNDIGATSYAAKDADDREKHPAEYTVDFDKKILKGELIKNQYVQKKND PKKPLTIYNITADLNGNRFTGSAKVNTEVKTRHADKEYLFFHTDADQRLEGGFFGDNG EELAGRFISNDNGVFGVFAGKQKTNASGTNPAMPFGKHTKILDSLKISVDEATDENPR PFEVSTMPDFGHPDKLLVEGREIPLVSKEKTIDLADGRKMTVSACCDFLTYVKLGRIK TERPAVKPKAQDEEDSGINNGEESEDEEEIAEESEDEVSEDDNGEDEDEIVEEEADEA EEIEEEAEEEEPEEESPEEGNGVSDGIPPAPEALKGRDIDLFLKGIRTAEADIPKTGT AHYTGTWEARIGEPIQWDNKADKAAKAEFDVDFGNKSISGTLTEQNGVEPAFRIENGV IEGNGFHPTARTRDNGINLSGNGSTNPQSFKADNLLVTGGFYGPQAAELGGTIFNKDG KSLGITEDIENEVENEADVGEQLEPEVKPQFGVVFGAKKDNKEVEK"
sig_peptide	310. .363
	/gene="lbpB"
	/note="signal peptidase II substrate"
mat_peptide	364. .2484
	/gene="lbpB"
	/product="lactoferrin binding protein B"
misc_feature	1902. .1911
	/gene="lbpB"
	/note="encodes gonococcal uptake sequence"

ORIGIN

Query Match 66.7%; Score 1416; DB 1; Length 2519;
Best Local Similarity 81.9%; Pred. No. 2.9e-286;
Matches 1732; Conservative 0; Mismatches 340; Indels 42; Gaps 7;





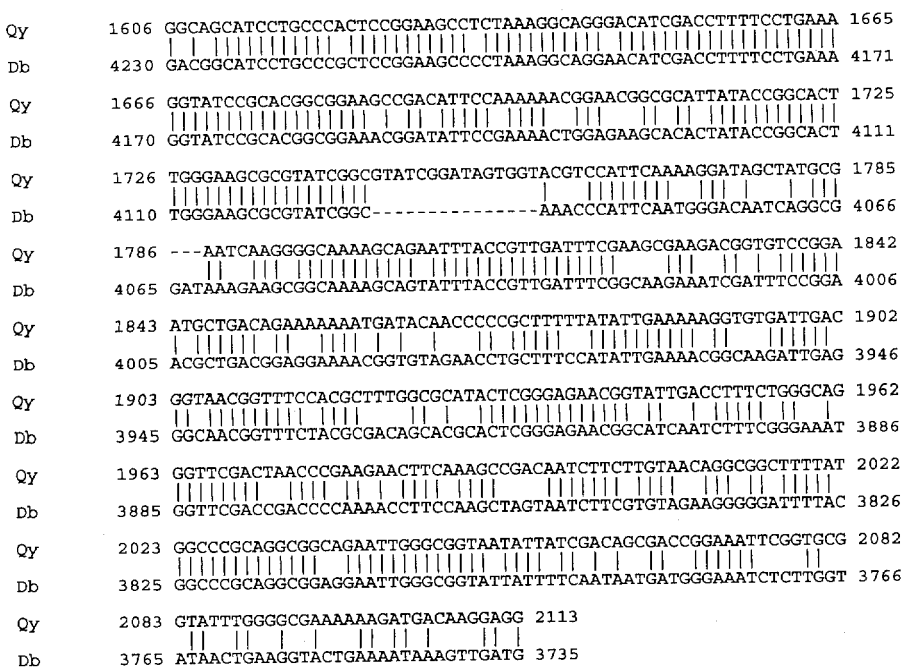


RESULT 12
AE002504/c
LOCUS AE002504 9955 bp DNA linear BCT 25-MAY-2000
DEFINITION Neisseria meningitidis serogroup B strain MC58 section 146 of 206 of the complete genome.

```
ACCESSION   AE002504 AE002098
VERSION     AE002504.1  GI:7226785
KEYWORDS    .
SOURCE      Neisseria meningitidis MC58
  ORGANISM  Neisseria meningitidis MC58
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE   1  (bases 1 to 9955)
  AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
  TITLE     Complete genome sequence of Neisseria meningitidis serogroup B
            strain MC58
  JOURNAL   Science 287 (5459), 1809-1815 (2000)
  MEDLINE   20175755
  PUBMED    10710307
REFERENCE   2  (bases 1 to 9955)
  AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .9955
                     /organism="Neisseria meningitidis MC58"
                     /mol_type="genomic DNA"
                     /strain="MC58"
                     /db_xref="taxon:122586"
                     /note="serogroup: B"
     gene            complement(791. .3622)
                     /gene="NMB1540"
     CDS             complement(791. .3622)
                     /gene="NMB1540"
                     /note="similar to PID:915278 percent identity: 97.88;
                     identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="lactoferrin-binding protein A"
                     /protein_id="AAF41895.1"
                     /db_xref="GI:7226786"
                     /translation="MNKKHGFPLTLTALAIATAFPAYAAQAGGATPDAAQTQSLKEIT
                     VRAAKVGRRSKEATGLGKIVKTSETLNKEQVLGIRDLTRYDPGVAVVEQGNGASGGYS
                     IRGVDKNRVAVSVDGVAQIQAFTVQGSLSGYGGRGGSGAINEIEYENISTVEIDKGAG
                     SSDHGSGALGGAVAFRTKEAADLISDGKSWGIQAKTAYGSKNRQFMKSLGAGFSKDGW
                     EGLLIRTERQGRETRPHGDIADGVEYGIDRLDAFRQTYDIKRKTREPFFSVEGERESK
                     PVAKLAGYGKYLNNQLNRWVKERIEQNQPLSAEEEAQVREAQARHENLSAQAYTGGGR
                     ILPDPMDYRSGSWLAKLGYRFGGRHYVGGVFEDTKQRYDIRDMTEKQYYGTDEAEKFR
                     DKSGVYDGDDFRDGLYFVPNIEEWKGDKNLVRGIGLKYSRTKFIDEHHRRRRMGLLYR
                     YENEAYSDNWADKAVLSFDKQGVATDNNTLKLNCAVYPAVDKSCRASADKPYSYDSSD
                     RFHYREQHNVLNASFEKSLKNKWTKHHLTLGFGYDASKAISRPEQLSHNAARISESTG
                     FDENNQDKYLLGKPEVVEGSVCGYIETLRSRKCVPRKINGSNIHISLNDRFSIGKYFD
                     FSLGGRYDRKNFTTSEELVRSGRYVDRSWNSGILFKPNRHFSVSYRASSGFRTPSFQE
                     LFGIDIYHDYPKGWQRPALKSEKAANREIGLQWKGDFGFLEISSFRNRYTDMIAVADH
                     KTKLPNQAGQLTEIDIRDYYNAQNMSLQGVNILGKIDWNGVYGKLPEGLYTTLAYNRI
                     KPKSVSNRPGLSLRSYALDAVQPSRYVLGFGYDQPEGKWGANIMLTYSKGKNPDELAY
                     LAGDQKRYSTKRASSSWSTADVSAYLNLKKRLTLRAAIYNIGNYRYVTWESLRQTAES
                     TANRHGGDSNYGRYAAPGRNFSLALEMKF"
     gene            complement(3619. .5832)
                     /gene="NMB1541"
     CDS             complement(3619. .5832)
                     /gene="NMB1541"
                     /note="similar to GP:3582728 percent identity: 86.78;
                     identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="lactoferrin-binding protein B"
                     /protein_id="AAF41896.1"
                     /db_xref="GI:7226787"
                     /translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
                     DVPTPPPAGSSVETTPVNRPAVGAAMRLPRRNIASYKQDGTEIPDKHQAEEHLPLKEK
                     DILFLDGTLKEQADKLKKKINERYSDVRVITSKKEEEKYQYQFVRAGYVFTRAEGKDN
                     EKEKTSDGKEFVNRFSYDGFVYYSGERPSQSLPSAGTVQYSGNWQYMTDAKRHRTGKA
                     VSSTDLGYTTYYGNEIGATSYEARDADDREKHPAEYTVDFDNKTLNGKLIKNQYVQNK
                     SNPNEPKKPLTIYDITATLDGNRFTGSAKVSTEVKTQHADKEYLFFHTDADQRLEGGF
                     FGDNGEELAGRFISNDNSVFGVFAGKQKTETENAADTKPALSSGKHTKILDSLKISVD
                     EASDKNPREFAISSMPDFGHPDKLLVEGREIPLVNKEQTIELADGRKTTIRTCCDFLT
                     YVKIGRMQTERPAAKPKAQDEERDEEDTGVDSVEEGEDEIDDEEGTEDAAVKDEGSEE
                     DEAVEGEDEAEEPEEESPTEEGGSGSDGILPAPEAPKGRNIDLFLKGIRTAETDIPKT
                     GEAHYTGTWEARIGKPIQWDNQADKEAAKAVFTVDFGKKSISGTLTEENGVEPAFHIE
                     NGKIEGNGFYATARTRENGINLSGNGSTDPKTFQASNLRVEGGFYGPQAEELGGIIFN
                     NDGKSLGITEGTENKVDVEAEVDAEVDVGKQLESEVKHQFGVVFGAKKDMQEVEK"
gene                 6855. .7076
                     /gene="NMB1542"
CDS                  6855. .7076
                     /gene="NMB1542"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF41897.1"
                     /db_xref="GI:7226788"
                     /translation="MSAVNVIFSCSYWRIFKKSDSCVLCFYQFRHGEPHKLINQENFS
                     KLYQAFDYIDSVGSNFPVIITTDGCGLFC"
gene                 7222. .8436
                     /gene="NMB1543"
CDS                  7222. .8436
                     /gene="NMB1543"
                     /note="conserved hypothetical protein; identified by
                     Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF41898.1"
                     /db_xref="GI:7226789"
                     /translation="MSEVEYFSHFISDGKGKLLEIPQRRGKQDGVFVDWISFTFHEDT
                     LLKVSGCPLFSDAEYMYVLSRKLEEILGFGITRKCKSRGNKFYESMYRLGSDDVDYGE
                     VHFGGQRNTVLVELKGTGCSVASPGWELRLKQFLDDSIRTRITRIDLALDFFDGEYTP
                     DQALLDHDNGFFDNSNQRPKSETIGTAWRNEDGSGKTFYVGRKKNSRFVRVYEKGRQL
                     GDKESKWVRFEIQFNYGDIEIPLDILINQGSYFCGAFPICRKFKNMPVPERFDQRKKK
                     LNLTFEHKLHYAKNAVGKLVNFMIEMGFDNSEIVESLKADSGFPKGLEPEKYALEMLR
                     DGLKHGFIHEQPDIDLEIELDELGVIAFKNSDKFDREKRLFSPDYDVEKERKYQEYLS
                     KVYHQNVDYDYF"
gene                 8448. .8759
                     /gene="NMB1544"
CDS                  8448. .8759
                     /gene="NMB1544"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF41899.1"
                     /db_xref="GI:7226790"
                     /translation="MFNQTQTVTYPATFLGAKKFKGEIDGSNIDTCSVLVATPLPAQS
                     GNAVGFTAAQMKFGDSKNFSKLENLKYPCEVMVTVEMTSTGKGMVPSLIDFQVAEKPK
                     G"
gene                 8763. .8963
                     /gene="NMB1545"
CDS                  8763. .8963
                     /gene="NMB1545"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
```

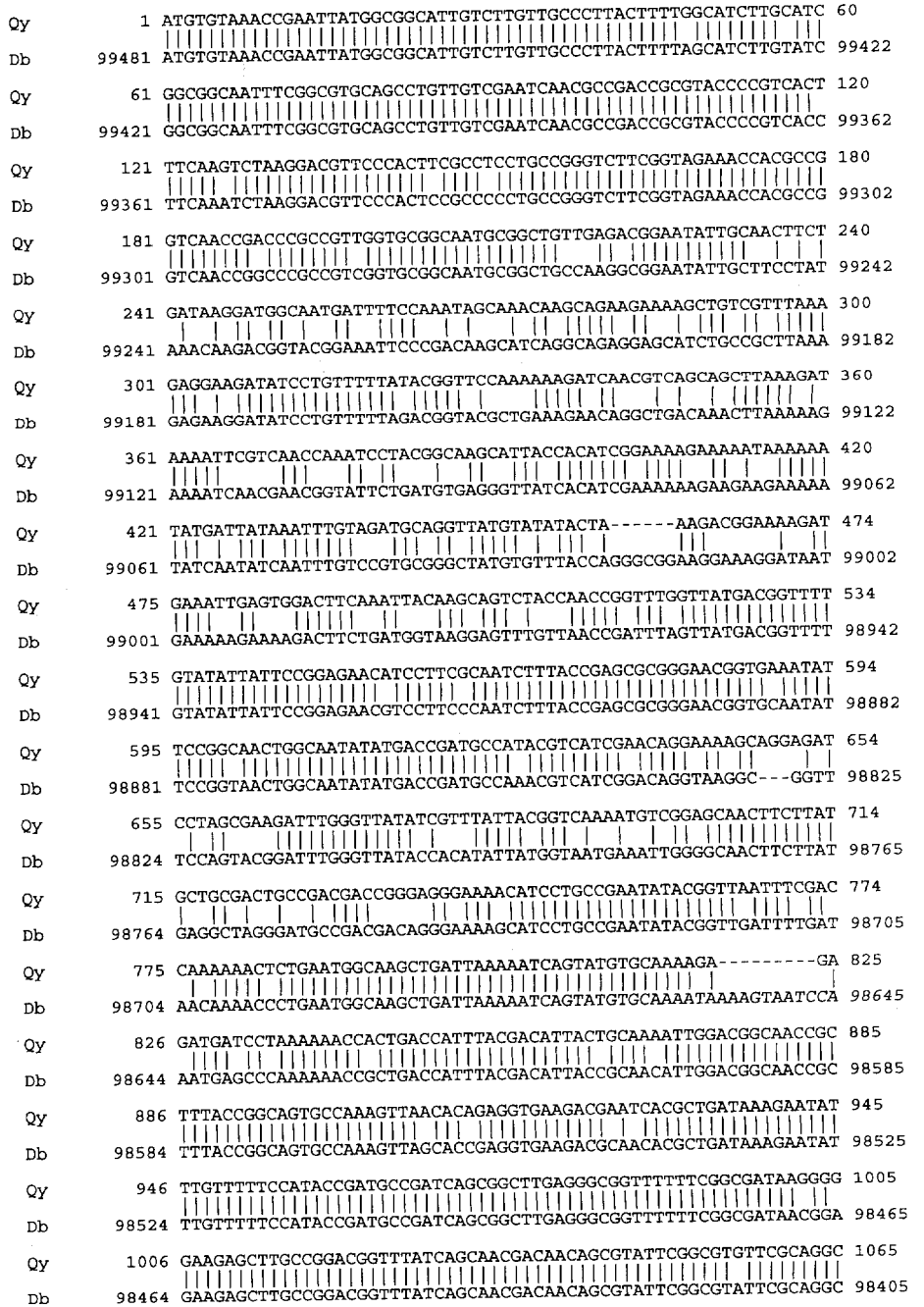
```
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF41900.1"
                     /db_xref="GI:7226791"
                     /translation="MKFEERFIVQDLETHDFIYPDPFGDVGFTQNIKSAGQFESYEDA
                     LNSGINEIGGGFQIFQFFVKSE"
     gene            9146. .9430
                     /gene="NMB1546"
     CDS             9146. .9430
                     /gene="NMB1546"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF41901.1"
                     /db_xref="GI:7226792"
                     /translation="MKQVKKSSYFKYQKRKKTMNIVKKYAVKAALAAGIFTPAIVMAD
                     TFDPSAIGTQVANVIMGFVSMVSAVGMAAITVILAIQGFKMAWSMIKSVK"
     gene            9458. .9760
                     /gene="NMB1547"
     CDS             9458. .9760
                     /gene="NMB1547"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF41902.1"
                     /db_xref="GI:7226793"
                     /translation="MGYRVGINCFDTRLQADDYLLSSLPPTVTQDGKIIRPERVGDKW
                     ILNGKPVTLSYPECSNFEQIKQGSYVGSTVLILFVVIYGFRLLINFLKDIGKVGTD"
ORIGIN

  Query Match           64.8%;  Score 1375.4;  DB 1;  Length 9955;
  Best Local Similarity 80.2%;  Pred. No. 1.1e-277;
  Matches 1709;  Conservative    0;  Mismatches  371;  Indels   51;  Gaps    6;

Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTAGCATCTTGTATC 5773

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACC 5713

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
              ||||| |||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db       5712 TTCAAATCTAAGGACGTTCCCACTCCGCCCCCTGCCGGGTCTTCGGTAGAAACCACGCCG 5653

Qy        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
              |||||||| |||||||| |||||||||||||||||  || ||||||||||  | | |
Db       5652 GTCAACCGGCCCGCCGTCGGTGCGGCAATGCGGCTGCCAAGGCGGAATATTGCTTCCTAT 5593

Qy        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
              | | || || |  ||  |||| | |   | || ||||| ||  | ||| || |||||
Db       5592 AAACAAGACGGTACGGAAATTCCCGACAAGCATCAGGCAGAGGAGCATCTGCCGCTTAAA 5533

Qy        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
              ||| | |||||||||||||||| ||||| |    ||||| ||   | |  | ||||| |
Db       5532 GAGAAGGATATCCTGTTTTTAGACGGTACGCTGAAAGAACAGGCTGACAAACTTAAAAAG 5473

Qy        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
              |||||     |||   || ||    |  ||  ||| |||||||| ||||  || |  |||||
Db       5472 AAAATCAACGAACGGTATTCTGATGTGAGGGTTATCACATCGAAAAAAGAAGAAGAAAAA 5413

Qy        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTA------AAGACGGAAAAGAT 474
              ||| | ||| |||||||    ||| || |||| | ||| |     |||   |  ||
Db       5412 TATCAATATCAATTTGTCCGTGCGGGCTATGTGTTTACCAGGGCGGAAGGAAAGGATAAT 5353

Qy        475 GAAATTGAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTT 534
              ||||  ||   ||||||  ||  ||| ||| |  ||||| ||| ||||||||||||||
```

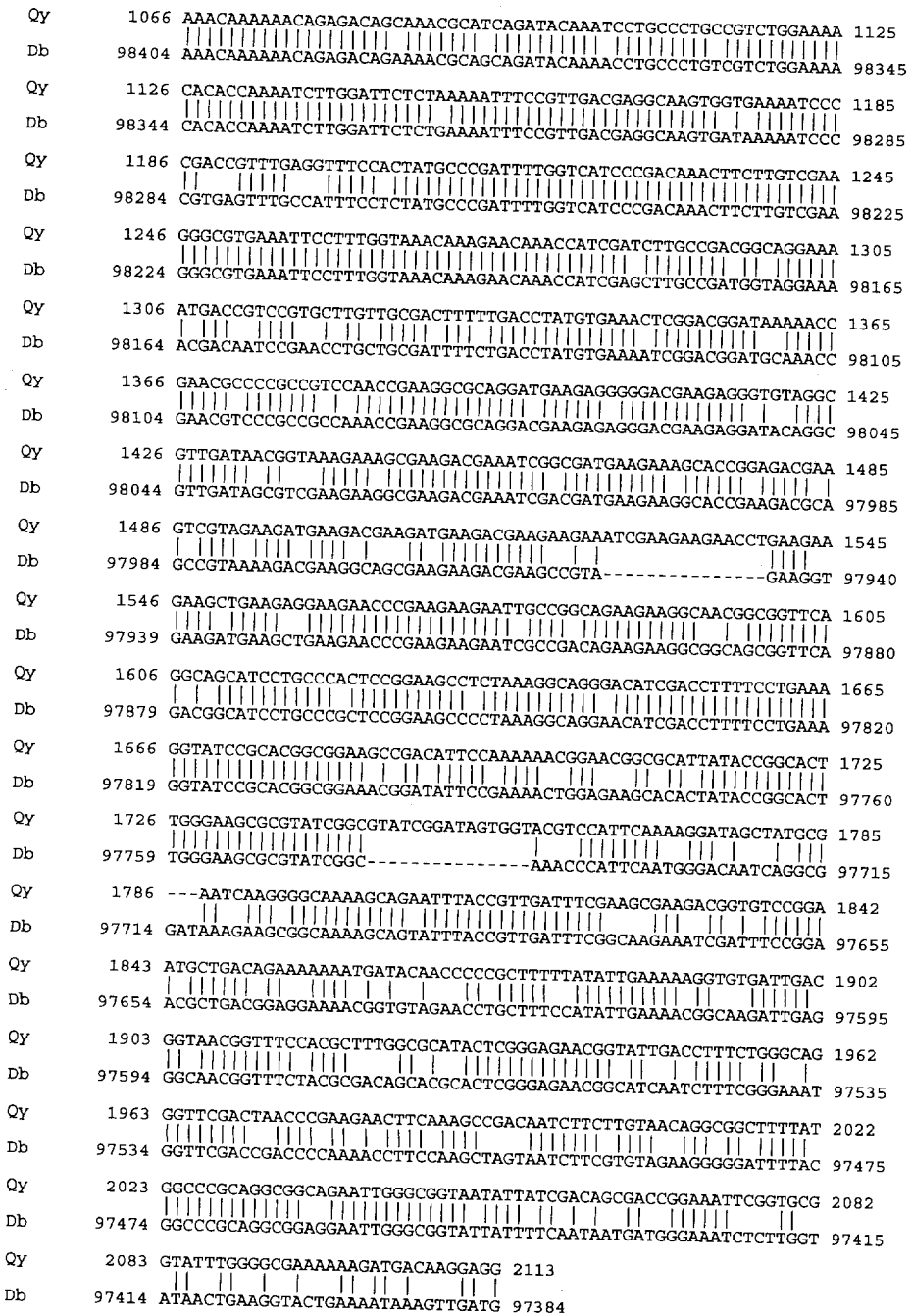
```
Db       5352 GAAAAAGAAAAGACTTCTGATGGTAAGGAGTTTGTTAACCGATTTAGTTATGACGGTTTT 5293

Qy         535 GTATATTATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATAT 594
               |||||||||||||||||||| |||||| |||||||||||||||||||||||||| |||||
Db       5292 GTATATTATTCCGGAGAACGTCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGCAATAT 5233

Qy         595 TCCGGCAACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGAT 654
               ||||| |||||||||||||||||||||| ||||||||| ||||| || ||    | |
Db       5232 TCCGGTAACTGGCAATATATGACCGATGCCAAACGTCATCGGACAGGTAAGGC---GGTT 5176

Qy         655 CCTAGCGAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTAT 714
               | ||    |||||||||||||| |   ||||| ||| |  |  | || ||||||||||||
Db       5175 TCCAGTACGGATTTGGGTTATACCACATATTATGGTAATGAAATTGGGGCAACTTCTTAT 5116

Qy         715 GCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGAC 774
               | || |  |  | |||| || ||| ||||||||||||||||||||| |||| ||
Db       5115 GAGGCTAGGGATGCCGACGACAGGGAAAAGCATCCTGCCGAATATACGGTTGATTTTGAT 5056

Qy         775 CAAAAAACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGA---------GA 825
               | ||||| |||||||||||||||||||||||||||||||||||||| |         |
Db       5055 AACAAAACCCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAATAAAAGTAATCCA 4996

Qy         826 GATGATCCTAAAAAACCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGC 885
               |||| || |||||||| |||||||||||||||||||| |||| |||||||||||||||
Db       4995 AATGAGCCCAAAAAACCGCTGACCATTTACGACATTACCGCAACATTGGACGGCAACCGC 4936

Qy         886 TTTACCGGCAGTGCCAAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATAT 945
               |||||||||||||||||||| ||| |||||||||||| | |||||||||||||||||||
Db       4935 TTTACCGGCAGTGCCAAAGTTAGCACCGAGGTGAAGACGCAACACGCTGATAAAGAATAT 4876

Qy         946 TTGTTTTTCCATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGG 1005
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       4875 TTGTTTTTCCATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAACGGA 4816

Qy        1006 GAAGAGCTTGCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGC 1065
               |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       4815 GAAGAGCTTGCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTATTCGCAGGC 4756

Qy        1066 AAACAAAAAACAGAGACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAA 1125
               |||||||||||| ||||||| |||||||| |||||||| |||||||||| ||||||||||
Db       4755 AAACAAAAAACAGAGACAGAAAACGCAGCAGATACAAAACCTGCCCTGTCGTCTGGAAAA 4696

Qy        1126 CACACCAAAATCTTGGATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCC 1185
               |||||||||||||||||||| |||||||||||||||||||||| | ||||||||
Db       4695 CACACCAAAATCTTGGATTCTCTGAAAATTTCCGTTGACGAGGCAAGTGATAAAATCCC 4636

Qy        1186 CGACCGTTTGAGGTTTCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAA 1245
               ||  |||||  ||||| |||||||||||||||||||||||||||||||||||||||||
Db       4635 CGTGAGTTTGCCATTTCCTCTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAA 4576

Qy        1246 GGGCGTGAAATTCCTTTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAA 1305
               |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       4575 GGGCGTGAAATTCCTTTGGTAAACAAAGAACAAACCATCGAGCTTGCCGATGGTAGGAAA 4516

Qy        1306 ATGACCGTCCGTGCTTGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACC 1365
               | |||  ||||  | || ||||| ||| |||||||||||||| |||||||||  |||||
Db       4515 ACGACAATCCGAACCTGCTGCGATTTTCTGACCTATGTGAAAATCGGACGGATGCAAACC 4456

Qy        1366 GAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGC 1425
               ||||| ||||||| | |||||||||||||| |||||| |||||||||||  |  ||||
Db       4455 GAACGTCCCGCCGCCAAACCGAAGGCGCAGGACGAAGAGAGGGACGAAGAGGATACAGGC 4396

Qy        1426 GTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAA 1485
               ||||||| ||  ||||| |||||||||||||||||||| |||||||||||| |||||| |
Db       4395 GTTGATAGCGTCGAAGAAGGCGAAGACGAAATCGACGATGAAGAAGGCACCGAAGACGCA 4336

Qy        1486 GTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAA 1545
               | |||| |||| |||| |   || |||||||||  | |                ||||
Db       4335 GCCGTAAAAGACGAAGGCAGCGAAGAAGACGAAGCCGTA--------------GAAGGT 4291

Qy        1546 GAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCA 1605
               |||| ||||| |||||||||||||||||| |||| |||||||||| | |||||||||
Db       4290 GAAGATGAAGCTGAAGAACCCGAAGAAGAATCGCCGACAGAAGAAGGCGGCAGCGGTTCA 4231
```



RESULT 13
AX044033/c
LOCUS AX044033 349980 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 112 from Patent WO0066791.
ACCESSION AX044033
VERSION AX044033.1 GI:11342917
KEYWORDS .
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE 1
AUTHORS Pizza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 112 09-NOV-2000;
CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES Location/Qualifiers
source 1. .349980
/organism="Neisseria meningitidis"
/mol_type="unassigned DNA"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.~seq 1: 1 to
349980 349980 bases~seq 108: 300001 to 649980 349980
bases~seq 109: 600001 to 949980 349980 bases~seq 110:
900001 to 1249980 349980 bases~seq 111: 1200001 to 1549980
349980 bases~seq 112: 1500001 to 1849980 349980 bases~seq
113: 1800001 to 2149980 349980 bases~seq 114: 2100001 to
2272325 172325 bases"
ORIGIN

Query Match 64.8%; Score 1375.4; DB 6; Length 349980;
Best Local Similarity 80.2%; Pred. No. 1.5e-277;
Matches 1709; Conservative 0; Mismatches 371; Indels 51; Gaps 6;



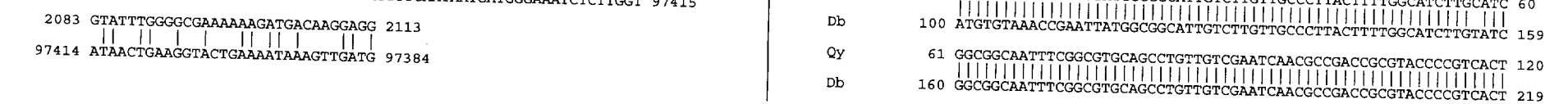


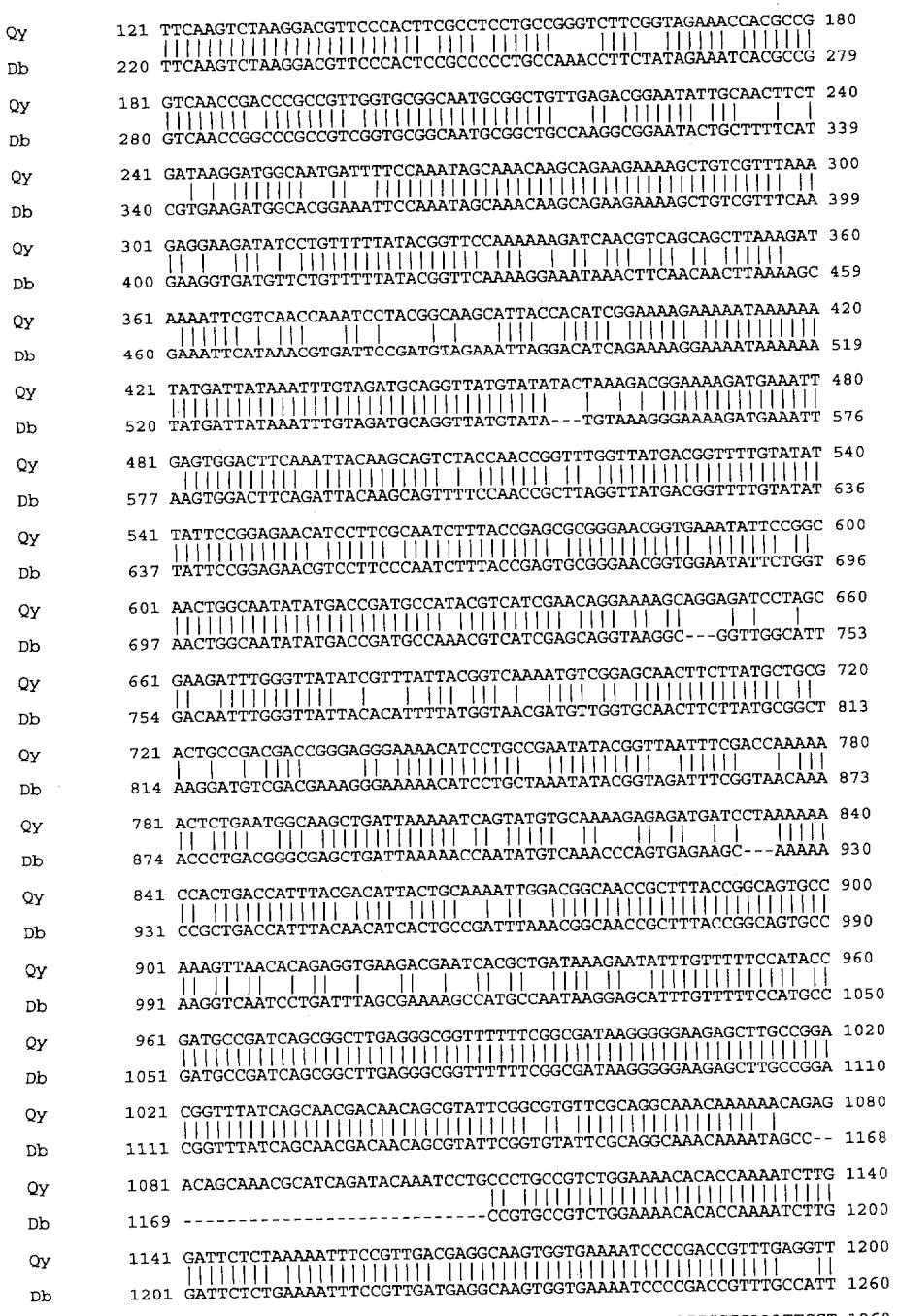
RESULT 14
AF022781
LOCUS AF022781 2277 bp DNA linear BCT 07-FEB-1998
DEFINITION Neisseria meningitidis lactoferrin binding protein B (lbpB) gene, complete cds.
ACCESSION AF022781
VERSION AF022781.1 GI:2843172
KEYWORDS .
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 2277)
AUTHORS Pettersson,A., Prinz,T., Umar,A., van der Biezen,J. and Tommassen,J.
TITLE Molecular characterization of LbpB, the second lactoferrin-binding protein of Neisseria meningitidis
JOURNAL Mol. Microbiol. 27 (3), 599-610 (1998)
MEDLINE 98149315
PUBMED 9489671
REFERENCE 2 (bases 1 to 2277)
AUTHORS Pettersson,A., Prinz,T., Umar,A., van der Biezen,J. and Tommassen,J.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1997) Department of Molecular Cell Biology, Utrecht University, Padualaan 8, Utrecht 3584 CH, The Netherlands

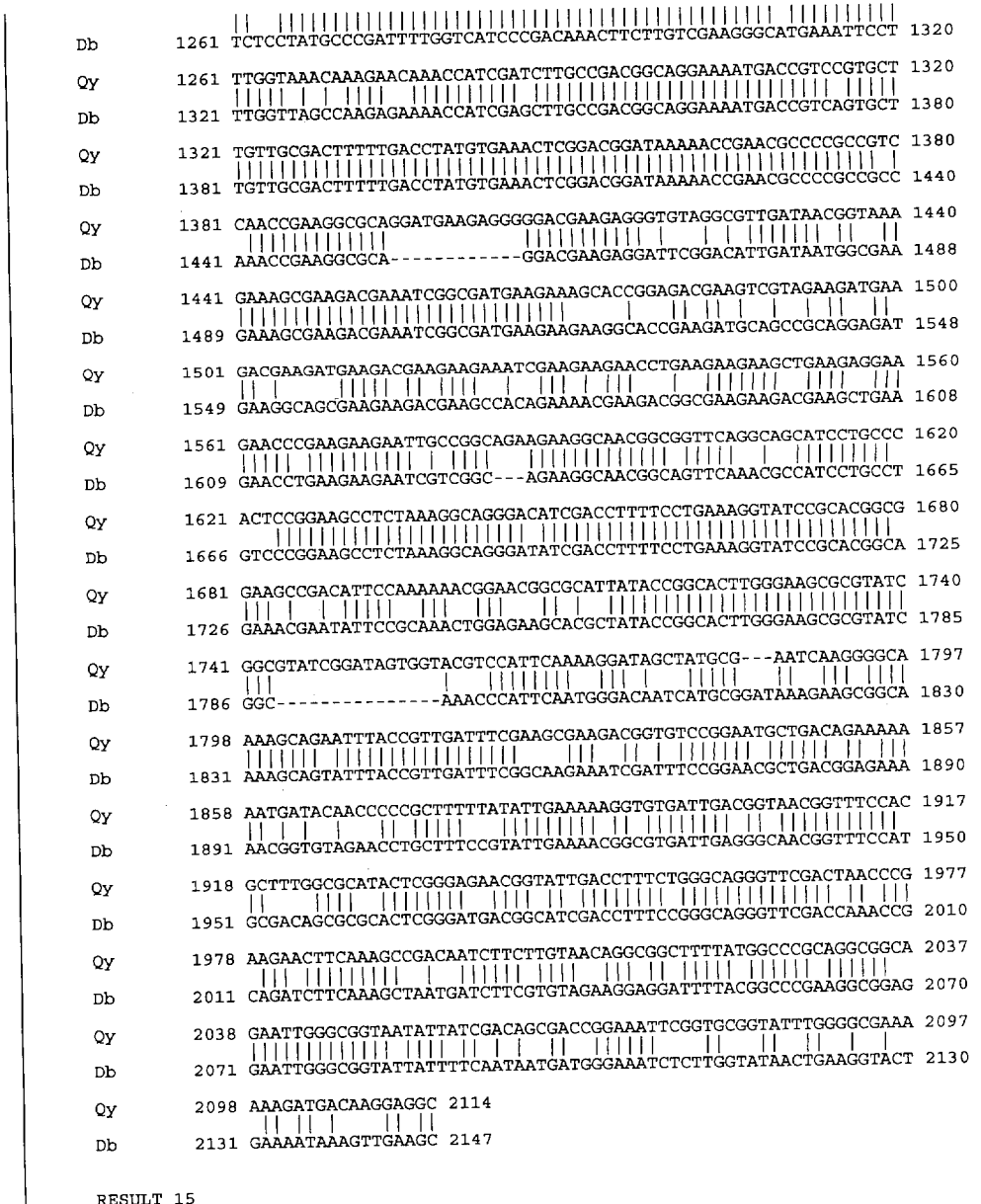
FEATURES Location/Qualifiers
source 1. .2277
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="BNCV"
/db_xref="taxon:487"
gene 57. .2277
/gene="lbpB"
misc_binding 57. .75
/gene="lbpB"
/note="putative binding site"
/bound_moiety="Fur"
CDS 100. .2277
/gene="lbpB"
/note="lactoferrin receptor"
/codon_start=1
/transl_table=11
/product="lactoferrin binding protein B"
/protein_id="AAC38143.1"
/db_xref="GI:2843173"
/translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
DVPTPPPAKPSIEITPVNRPAVGAAMRLPRRNTAFHREDGTEIPNSKQAEEKLSFQEG
DVLFLYGSKGNKLQQLKSEIHKRDSDVEIRTSEKENKKYDYKFVDAGYVYVKGKDEIK
WTSDYKQFSNRLGYDGFVYYSGERPSQSLPSAGTVEYSGNWQYMTDAKRHRAGKAVGI
DNLGYYTFYGNDVGATSYAAKDVDEREKHPAKYTVDFGNKTLTGELIKNQYVKPSEKQ
KPLTIYNITADLNGNRFTGSAKVNPDLAKSHANKEHLFFHADADQRLEGGFFGDKGEE
LAGRFISNDNSVFGVFAGKQNSPVPSGKHTKILDSLKISVDEASGENPRPFAISPMPD
FGHPDKLLVEGHEIPLVSQEKTIELADGRKMTVSACCDFLTYVKLGRIKTERPAAKPK
AQDEEDSDIDNGEESEDEIGDEEEGTEDAAAGDEGSEEDEATENEDGEEDEAEEPEEE
SSAEGNGSSNAILPVPEASKGRDIDLFLKGIRTAETNIPQTGEARYTGTWEARIGKPI
QWDNHADKEAAKAVFTVDFGKKSISGTLTEKNGVEPAFRIENGVIEGNGFHATARTRD
DGIDLSGQGSTKPQIFKANDLRVEGGFYGPKAEELGGIIFNNDGKSLGITEGTENKVE
ADVDVDVDVDVDADADVEQLKPEVKPQFGVVFGAKKDNKEVEK"

ORIGIN

Query Match 61.0%; Score 1296.2; DB 1; Length 2277;
Best Local Similarity 79.0%; Pred. No. 3.8e-261;
Matches 1672; Conservative 0; Mismatches 373; Indels 72; Gaps 8;







RESULT 15
A98968
LOCUS A98968 2277 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 1 from Patent WO9909176.
ACCESSION A98968
VERSION A98968.1 GI:6781928
KEYWORDS .
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis

```
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE   1  (bases 1 to 2277)
  AUTHORS   Pettersson-Fernholm,A.M. and Tommassen,J.P.
  TITLE     NEISSERIA LACTOFERRIN BINDING PROTEIN
  JOURNAL   Patent: WO 9909176-A 1 25-FEB-1999;
            UNIV UTRECHT (NL); PETTERSSON FERNHOLM ANNIKA MAR (NL)
FEATURES             Location/Qualifiers
     source          1. .2277
                     /organism="Neisseria meningitidis"
                     /mol_type="unassigned DNA"
                     /strain="BNCV"
                     /db_xref="taxon:487"
     CDS             100. .2277
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB70581.1"
                     /db_xref="GI:6781929"
                     /db_xref="REMTREMBL:CAB70581"
                     /translation="MCKPNYGGIVLLPLLLASCIGGNFGVQPVVESTPTAYPVTFKSK
                     DVPTPPPAKPSIEITPVNRPAVGAAMRLPRRNTAFHREDGTEIPNSKQAEEKLSFQEG
                     DVLFLYGSKGNKLQQLKSEIHKRDSDVEIRTSEKENKKYDYKFVDAGYVYVKGKDEIK
                     WTSDYKQFSNRLGYDGFVYYSGERPSQSLPSAGTVEYSGNWQYMTDAKRHRAGKAVGI
                     DNLGYYTFYGNDVGATSYAAKDVDEREKHPAKYTVDFGNKTLTGELIKNQYVKPSEKQ
                     KPLTIYNITADLNGNRFTGSAKVNPDLAKSHANKEHLFFHADADQRLEGGFFGDKGEE
                     LAGRFISNDNSVFGVFAGKQNSPVPSGKHTKILDSLKISVDEASGENPRPFAISPMPD
                     FGHPDKLLVEGHEIPLVSQEKTIELADGRKMTVSACCDFLTYVKLGRIKTERPAAKPK
                     AQDEEDSDIDNGEESEDEIGDEEEGTEDAAAGDEGSEEDEATENEDGEEDEAEEPEEE
                     SSAEGNGSSNAILPVPEASKGRDIDLFLKGIRTAETNIPQTGEARYTGTWEARIGKPI
                     QWDNHADKEAAKAVFTVDFGKKSISGTLTEKNGVEPAFRIENGVIEGNGFHATARTRD
                     DGIDLSGQGSTKPQIFKANDLRVEGGFYGPKAEELGGIIFNNDGKSLGITEGTENKVE
                     ADVDVDVDVDVDADADVEQLKPEVKPQFGVVFGAKKDNKEVEK"
ORIGIN

  Query Match          61.0%;  Score 1296.2;  DB 6;  Length 2277;
  Best Local Similarity  79.0%;  Pred. No. 3.8e-261;
  Matches 1672;  Conservative    0;  Mismatches  373;  Indels   72;  Gaps    8;

Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        100 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGTATC 159

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 219

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
              |||||||||||||||||||||| |||| |||||  |||| ||||||
Db        220 TTCAAGTCTAAGGACGTTCCCACTCCGCCCCCTGCCAAACCTTCTATAGAAATCACGCCG 279

Qy        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
              |||||||| |||||||| ||||||||||||||||||  || |||||||  |||  | |
Db        280 GTCAACCGGCCCGCCGTCGGTGCGGCAATGCGGCTGCCAAGGCGGAATACTGCTTTTCAT 339

Qy        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
              | | ||||||| ||  ||||||||||||||||||||||||||||||||||||||||||||
Db        340 CGTGAAGATGGCACGGAAATTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTCAA 399

Qy        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
              || |  ||| | |||||||||||||||||| |||  | || ||| ||| || ||||||
Db        400 GAAGGTGATGTTCTGTTTTTATACGGTTCAAAAGGAAATAAACTTCAACAACTTAAAAGC 459

Qy        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
              ||||||  |||  || |   | |  |||| ||||| |||||| ||||||||||||||
Db        460 GAAATTCATAAACGTGATTCCGATGTAGAAATTAGGACATCAGAAAAGGAAAATAAAAAA 519

Qy        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
              |||||||||||||||||||||||||||||||||||  | | | ||||||||||||||||
Db        520 TATGATTATAAATTTGTAGATGCAGGTTATGTATA---TGTAAAGGGAAAAGATGAAATT 576

Qy        481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
              |||||||||||  |||||||||||  | ||||||| || ||||||||||||||||||||||

Db        577 AAGTGGACTTCAGATTACAAGCAGTTTTCCAACCGCTTAGGTTATGACGGTTTTGTATAT 636

Qy        541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
              |||||||||||||| |||||| |||||||||||||| |||||||||||| |||||||||
Db        637 TATTCCGGAGAACGTCCTTCCCAATCTTTACCGAGTGCGGGAACGGTGGAATATTCTGGT 696

Qy        601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
              |||||||||||||||||||||||||| |||||||||| |||| || ||    | | |
Db        697 AACTGGCAATATATGACCGATGCCAAACGTCATCGAGCAGGTAAGGC---GGTTGGCATT 753

Qy        661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
              ||  |||||||||| |  | ||| ||| |  |||| || ||||||||||||||| ||
Db        754 GACAATTTGGGTTATTACACATTTTATGGTAACGATGTTGGTGCAACTTCTTATGCGGCT 813

Qy        721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
              | |  | ||||      || ||||||||||||  ||||||||||  ||||||  | |||
Db        814 AAGGATGTCGACGAAAGGGAAAAACATCCTGCTAAATATACGGTAGATTTCGGTAACAAA 873

Qy        781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
              || ||||  ||| |||||||||||||| || |||||  ||  ||  || |  |||||
Db        874 ACCCTGACGGGCGAGCTGATTAAAAACCAATATGTCAAACCCAGTGAGAAGC---AAAAA 930

Qy        841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
              || |||||||||||  |||| |||||  | ||  |||||||||||||||||||||||||
Db        931 CCGCTGACCATTTACAACATCACTGCCGATTTAAACGGCAACCGCTTTACCGGCAGTGCC 990

Qy        901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
              || || ||  | || |   ||  | || ||  |||| ||  ||||||||||||||  ||
Db        991 AAGGTCAATCCTGATTTAGCGAAAAGCCATGCCAATAAGGAGCATTTGTTTTTCCATGCC 1050

Qy        961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1051 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1110

Qy       1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
              |||||||||||||||||||||||||||||||||| || ||||||||||||||||| |
Db       1111 CGGTTTATCAGCAACGACAACAGCGTATTCGGTGTATTCGCAGGCAAACAAAATAGCC-- 1168

Qy       1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 ----------------------------CCGTGCCGTCTGGAAAACACACCAAAATCTTG 1200

Qy       1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
              |||||||| |||||||||||||| ||||||||||||||||||||||||||||||||  ||
Db       1201 GATTCTCTGAAAATTTCCGTTGATGAGGCAAGTGGTGAAAATCCCCGACCGTTTGCCATT 1260

Qy       1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCTCCTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCATGAAATTCCT 1320

Qy       1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
              ||||| | |||  |||||||||| |||||||||||||||||||||||||||||||||||
Db       1321 TTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCGTCAGTGCT 1380

Qy       1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1381 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGCC 1440

Qy       1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
              |||||||||||||  |||||||||| |   | |||||||  ||  ||
Db       1441 AAACCGAAGGCGCA----------GGACGAAGAGGATTCGGACATTGATAATGGCGAA 1488

Qy       1441 GAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAA 1500
              |||||||||||||||||||||||||||||||||||  |  || || |  |  || ||
Db       1489 GAAAGCGAAGACGAAATCGGCGATGAAGAAGAAGGCACCGAAGATGCAGCCGCAGGAGAT 1548

Qy       1501 GACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAA 1560
              || |  |||| || ||||  | ||| | ||| |  |||||||  |||| |||
Db       1549 GAAGGCAGCGAAGAAGACGAAGCCACAGAAAACGAAGACGGCGAAGAAGACGAAGCTGAA 1608

Qy       1561 GAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCC 1620
              |||||  |||||||||| | ||||   |||||||||||||| |||||  |  |||||||||
Db       1609 GAACCTGAAGAAGAATCGTCGGC---AGAAGGCAACGGCAGTTCAAACGCCATCCTGCCT 1665
```

```
Qy       1621 ACTCCGGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCG 1680
Db       1666 GTCCCGGAAGCCTCTAAAGGCAGGGATATCGACCTTTTCCTGAAAGGTATCCGCACGGCA 1725

Qy       1681 GAAGCCGACATTCCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATC 1740
Db       1726 GAAACGAATATTCCGCAAACTGGAGAAGCACGCTATACCGGCACTTGGGAAGCGCGTATC 1785

Qy       1741 GGCGTATCGGATAGTGGTACGTCCATTCAAAAGGATAGCTATGCG---AATCAAGGGGCA 1797
Db       1786 GGC---------------AAACCCATTCAATGGGACAATCATGCGGATAAAGAAGCGGCA 1830

Qy       1798 AAAGCAGAATTTACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAA 1857
Db       1831 AAAGCAGTATTTACCGTTGATTTCGGCAAGAAATCGATTTCCGGAACGCTGACGGAGAAA 1890

Qy       1858 AATGATACAACCCCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCAC 1917
Db       1891 AACGGTGTAGAACCTGCTTTCCGTATTGAAAACGGCGTGATTGAGGGCAACGGTTTCCAT 1950

Qy       1918 GCTTTGGCGCATACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCG 1977
Db       1951 GCGACAGCGCGCACTCGGGATGACGGCATCGACCTTTCCGGGCAGGGTTCGACCAAACCG 2010

Qy       1978 AAGAACTTCAAAGCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCA 2037
Db       2011 CAGATCTTCAAAGCTAATGATCTTCGTGTAGAAGGAGGATTTTACGGCCCGAAGGCGGAG 2070

Qy       2038 GAATTGGGCGGTAATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAA 2097
Db       2071 GAATTGGGCGGTATTATTTTCAATAATGATGGGAAATCTCTTGGTATAACTGAAGGTACT 2130

Qy       2098 AAAGATGACAAGGAGGC 2114
Db       2131 GAAAATAAAGTTGAAGC 2147
```

Search completed: August 24, 2004, 23:48:01
Job time : 7978.58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 08:48:44 ; Search time 719.611 Seconds
(without alignments)
12538.967 Million cell updates/sec

Title: US-10-735-098-9
Perfect score: 2124
Sequence: 1 atgtgtaaaccgaattatgg.....acaaggaggcaacacgatga 2124

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2124	100.0	2124	2	AAX23323	Aax23323 N. mening
	2	1562.4	73.6	2169	2	AAX23320	Aax23320 N. mening
	3	1516.8	71.4	2226	7	ACA41945	Aca41945 Prokaryot
	4	1499.8	70.6	2226	2	AAX23321	Aax23321 N. mening
c	5	1375.4	64.8	110000	3	AAA81489_1	Continuation (2 of
c	6	1375.4	64.8	349980	3	AAF21611	Aaf21611 Neisseria
	7	1296.2	61.0	2277	2	AAX23319	Aax23319 N. mening
	8	1291.8	60.8	2262	2	AAX23322	Aax23322 N. mening
c	9	937.4	44.1	14652	3	AAA81482	Aaa81482 N. mening
	10	497.6	23.4	1000	4	AAF91389	Aaf91389 N. mening
	11	497.6	23.4	1000	6	ABK37769	Abk37769 DNA seque
	12	428.8	20.2	707	3	AAA81815	Aaa81815 N. mening
	13	104.6	4.9	3300	6	ABS67377	Abs67377 Neisseria
	14	98.2	4.6	708	5	AAS69547	Aas69547 DNA encod
	15	98.2	4.6	708	5	AAS75461	Aas75461 DNA encod
	16	96.8	4.6	801	5	AAS90729	Aas90729 DNA encod
	17	95.6	4.5	2334	5	AAS90705	Aas90705 DNA encod
	18	95	4.5	400	5	AAS75460	Aas75460 DNA encod
	19	95	4.5	963	5	AAS68580	Aas68580 DNA encod
	20	94.4	4.4	49999	2	AAZ23896	Aaz23896 Murine LO
	21	94.4	4.4	49999	2	AAZ23891	Aaz23891 Murine LO
c	22	94.2	4.4	453	5	AAS92079	Aas92079 DNA encod
	23	93.8	4.4	248	5	AAS90707	Aas90707 DNA encod
c	24	93.8	4.4	305	4	AAI21797	Aai21797 Probe #11
c	25	93.8	4.4	305	4	ABA66871	Aba66871 Human foe
c	26	93.8	4.4	305	4	AAI47083	Aai47083 Probe #15
c	27	93.8	4.4	305	4	ABA48950	Aba48950 Human bre
c	28	93.8	4.4	305	4	ABA33942	Aba33942 Probe #12
c	29	93.8	4.4	305	4	AAK41027	Aak41027 Human bon
c	30	93.8	4.4	305	4	AAK15306	Aak15306 Human bra
c	31	93.8	4.4	305	4	ABS40624	Abs40624 Human liv
c	32	93.8	4.4	305	5	AAI07480	Aai07480 Probe #74
c	33	93.8	4.4	305	6	ABS15007	Abs15007 Human gen
	34	93.8	4.4	483	5	AAS69549	Aas69549 DNA encod
c	35	93.8	4.4	496	4	AAI12609	Aai12609 Probe #25
c	36	93.8	4.4	496	4	ABA54309	Aba54309 Human foe
c	37	93.8	4.4	496	4	AAI33960	Aai33960 Probe #26
c	38	93.8	4.4	496	4	ABA43847	Aba43847 Human bre
c	39	93.8	4.4	496	4	ABA24068	Aba24068 Probe #25
c	40	93.8	4.4	496	4	AAK28026	Aak28026 Human bon
c	41	93.8	4.4	496	4	AAK02592	Aak02592 Human bra
c	42	93.8	4.4	496	4	ABS27626	Abs27626 Human liv
c	43	93.8	4.4	496	5	AAI02512	Aai02512 Probe #25
c	44	93.8	4.4	496	6	ABS02506	Abs02506 Human gen
	45	93.2	4.4	390	5	AAS69541	Aas69541 DNA encod

ALIGNMENTS

RESULT 1
AAX23323
ID AAX23323 standard; cDNA; 2124 BP.
XX
AC AAX23323;
XX
DT 11-JUN-1999 (first entry)
XX
DE N. meningitidis strain 881607 LbpB cDNA.
XX
KW LbpB; lactoferrin binding protein; vaccine; neisserial disease;
KW meningitis; diagnosis; treatment; ds.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1. .2124
FT /*tag= a
FT /product= "LbpB"
XX
PN WO9909176-A1.
XX
PD 25-FEB-1999.
XX
PF 10-AUG-1998; 98WO-EP005117.
XX
PR 15-AUG-1997; 97GB-00017423.
PR 05-FEB-1998; 98GB-00002544.
XX
PA (UYUT-) RIJKSUNIV UTRECHT.
PA (TECH-) TECHNOLOGY FOUND TECHNOLOGIESTICHTING ST.
XX
PI Pettersson-Fernholm AM, Tommassen JPM;
XX
DR WPI; 1999-190165/16.
DR P-PSDB; AAW93496.
XX
PT New lactoferrin-binding protein B polynucleotides - obtained from
PT Neisseria meningitidis, used to develop products for the diagnosis,
PT prevention and treatment of neisserial disease, e.g. meningitis.
XX
PS Claim 2; Page 105-109; 116pp; English.
XX
CC This invention describes novel lactoferrin-binding protein B (LbpB)
CC strains of Neisseria meningitidis. The products of this invention can be
CC used for vaccinating humans against neisserial disease e.g. meningitis.

CC Antibodies raised against the proteins of the invention can be used for
CC diagnosing or treating neisserial disease in humans. The LbpB
CC polypeptides can also be used for identifying compounds which inhibit the
CC polypeptides
XX
SQ Sequence 2124 BP; 663 A; 461 C; 541 G; 459 T; 0 U; 0 Other;

Query Match 100.0%; Score 2124; DB 2; Length 2124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60

Qy   61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy  121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180

Qy  181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240

Qy  241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300

Qy  301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360

Qy  361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420

Qy  421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480

Qy  481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540

Qy  541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600

Qy  601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660

Qy  661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720

Qy  721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780

Qy  781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840

Db  781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGGACGGCAACCGCTTTACCGGCAGTGCC 900

Qy  841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900

Qy  901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db  901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960

Qy  961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020

Qy 1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080

Qy 1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140

Qy 1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200

Qy 1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260

Qy 1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320

Qy 1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380

Qy 1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440

Qy 1441 GAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAA 1500

Qy 1501 GACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 GACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAA 1560

Qy 1561 GAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 GAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCC 1620

Qy 1621 ACTCCGGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 ACTCCGGAAGCCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCG 1680

Qy 1681 GAAGCCGACATTCCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 GAAGCCGACATTCCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATC 1740

Qy 1741 GGCGTATCGGATAGTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 GGCGTATCGGATAGTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAA 1800

Qy 1801 GCAGAATTTACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAAT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 GCAGAATTTACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAAT 1860

Qy 1861 GATACAACCCCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 GATACAACCCCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCT 1920

Qy 1921 TTGGCGCATACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 TTGGCGCATACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAG 1980

Qy 1981 AACTTCAAAGCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAA 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 AACTTCAAAGCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAA 2040
```

```
Qy       2041 TTGGGCGGTAATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAA 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTGGGCGGTAATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAA 2100

Qy       2101 GATGACAAGGAGGCAACACGATGA 2124
             |||||||||||||||||||||||||
Db       2101 GATGACAAGGAGGCAACACGATGA 2124
```

RESULT 2
AAX23320
ID AAX23320 standard; cDNA; 2169 BP.
XX
AC AAX23320;
XX
DT 11-JUN-1999 (first entry)
XX
DE N. meningitidis strain M981 LbpB cDNA.
XX
KW LbpB; lactoferrin binding protein; vaccine; neisserial disease;
KW meningitis; diagnosis; treatment; ds.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1. .2169
FT /*tag= a
FT /product= "LbpB"
XX
PN WO9909176-A1.
XX
PD 25-FEB-1999.
XX
PF 10-AUG-1998; 98WO-EP005117.
XX
PR 15-AUG-1997; 97GB-00017423.
PR 05-FEB-1998; 98GB-00002544.
XX
PA (UYUT-) RIJKSUNIV UTRECHT.
PA (TECH-) TECHNOLOGY FOUND TECHNOLOGIESTICHTING ST.
XX
PI Pettersson-Fernholm AM, Tommassen JPM;
XX
DR WPI; 1999-190165/16.
DR P-PSDB; AAW93493.
XX
PT New lactoferrin-binding protein B polynucleotides - obtained from
PT Neisseria meningitidis, used to develop products for the diagnosis,
PT prevention and treatment of neisserial disease, e.g. meningitis.
XX
PS Claim 2; Page 82-86; 116pp; English.
XX
CC This invention describes novel lactoferrin-binding protein B (LbpB)
CC strains of Neisseria meningitidis. The products of this invention can be
CC used for vaccinating humans against neisserial disease e.g. meningitis.
CC Antibodies raised against the proteins of the invention can be used for
CC diagnosing or treating neisserial disease in humans. The LbpB
CC polypeptides can also be used for identifying compounds which inhibit the
CC polypeptides
XX
SQ Sequence 2169 BP; 675 A; 457 C; 550 G; 487 T; 0 U; 0 Other;

Query Match 73.6%; Score 1562.4; DB 2; Length 2169;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1835; Conservative 0; Mismatches 286; Indels 51; Gaps 5;

```
Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCCCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180

Qy        181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
              ||||||| ||||||||||||||||||||||||||||||||||||||||||| |||   | |
Db        181 GTCAACCAGCCCGCCGTCGGTGCGGCAATGCGGCTGTTGAGACGGAATACTGCTTTTCAT 240

Qy        241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
              | | ||||||| | |||| ||||||||||||||||||||||||||||||||||||||
Db        241 CGTGAAGATGGCACGGCAATTCCCGATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300

Qy        301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
              || | ||| | |||||||||||||||||| ||| || || ||| ||| || ||||||
Db        301 GAAGGTGATGTTCTGTTTTTATACGGTTCAAAAGAAAATAAACTTCAACAACTTAAAAGC 360

Qy        361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
              |||||| | ||| |||||| ||||||||||||||||||||||| ||||||||||
Db        361 GAAATTCATAAACGTAATCCTGAGGCAAGCATTACCACATCGGAAAATGAAAATAAAAAA 420

Qy        421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
              ||| |||||   ||||| ||| |||||||| | ||||||| ||||||||||||||||||
Db        421 TATAATTATCGGTTTGTCAGTGCCGGTTATGTGTTTACTAAAAACGGAAAAGATGAAATT 480

Qy        481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
              ||| || || || | | || || || || |||||||||||||||||||||||||||||||
Db        481 GAGAAAACATCGGATGAAAAGCAGTTTTCTAATCGTTTAGGCTATGACGGTTTTGTATAT 540

Qy        541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
              ||| ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        541 TATCTCGGAGAACATCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600

Qy        601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
              |||||||||||||||||||||||||||||||||| | ||| ||   || | | | ||
Db        601 AACTGGCAATATATGACCGATGCCATACGTCATCGGAGAGGTAA---GGGGGTTTCCAGT 657

Qy        661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
              | |||||||||||| | ||||| ||| | | | || ||| |||||||||| ||
Db        658 GTGGATTTGGGTTATACCACATATTATGGTAATGAAATTGGGGCAGCTTCTTATGAGGCT 717

Qy        721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
              | | | ||| | || |||||||||||||||||||||||||||||||||| |||||
Db        718 AGGGATGCCGATGGCCGGGAAAAACATCCTGCCGAATATACGGTTAATTTCGACAAAAAA 777

Qy        781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
              | ||| | || ||| ||||||||||||||||||||||||||||||||||||||||||||
Db        778 AACCTGGAAGGTAAGTTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAT 837

Qy        841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
              |||||||||||||| |||||| |||| ||||||||||||||||||||||||||||||||
Db        838 CCACTGACCATTTACAACATTACCGCAACATTGGACGGCAACCGCTTTACCGGCAGTGCC 897

Qy        901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
              ||||||| ||| |||||||||||| | ||||||||||||||||||||||||||||||||
Db        898 AAAGTTAGCACCGAGGTGAAGACGCAACACGCTGATAAAGAATATTTGTTTTTCCATACC 957

Qy        961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
              |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db        958 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAACGGAGAAGAGCTTGCCGGG 1017

Qy       1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 CGGTTTATCAGTAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1077

Qy       1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1137

Qy       1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
              |||||||||||||||||||||||||||||| | || ||| || |||| ||||| ||
Db       1138 GATTCTCTAAAAATTTCCGTTGACGAGGCGACTGATGACCATGCCCGTAAGTTTGCCATT 1197
```

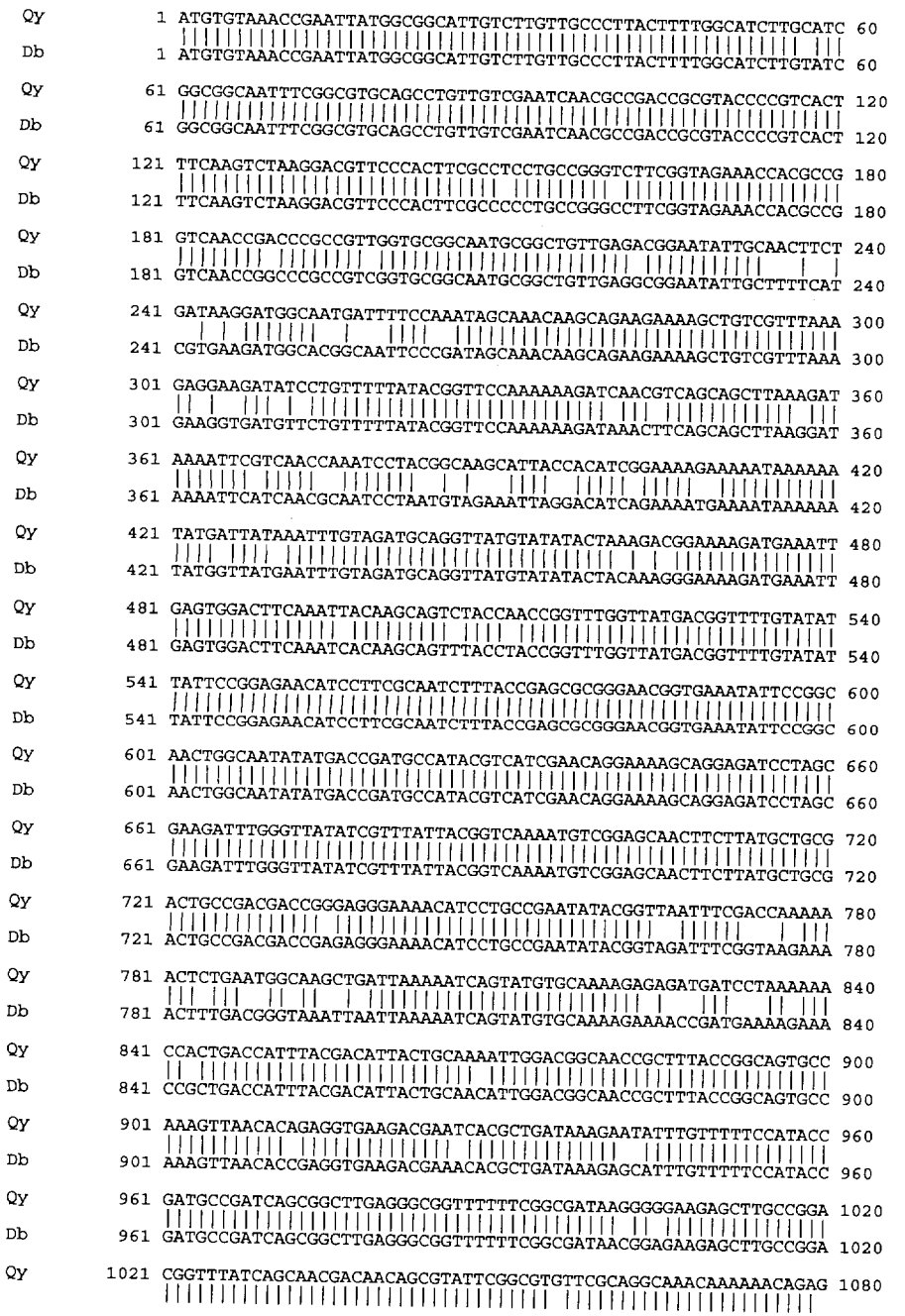
```
Qy       1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1257

Qy       1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
             |||||  |  | ||||  |||||||||| |||||||||||||||||||||||| ||||||||
Db       1258 TTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCATCCGTGCT 1317

Qy       1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
             |||||||| ||| ||||||||||||||||||||||||||||||||| ||||||||||||
Db       1318 TGTTGCGATTTTCTGACCTATGTGAAACTCGGACGGATAAAAACCGACCGCCCCGCCGTC 1377

Qy       1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
             ||||||||||||||||||||||||||     ||    |  ||||  || | ||    |
Db       1378 AAACCGAAGGCGCAGGATGAAGAGGATTCGGACATTGATAATGGCGAAGAAAGCGAAGAC 1437

Qy       1441 GAAA---GCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGA---------------- 1481
             ||||    |||||| || | |||||| || |||  |||||  |||
Db       1438 GAAATTTCCGAAGATGATAACGGCGAAGATGAAGTCACCGAAGAAGAGGAAGCTGAAGAA 1497

Qy       1482 --CGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAA------GAA 1533
                |||||  | |   ||||||||||||||| |||||     |||||||  |||       |||
Db       1498 ACCGAAGAAGAAACTGATGAAGACGAAGAGGAAGAACCCGAAGAAACTGAAGAAACTGAA 1557

Qy       1534 GAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGC 1593
             |||  ||||||||     |||   |||||||  | ||||||  |||  |||| ||||||||||
Db       1558 GAAACTGAAGAAACTGAAGAAACTGAAGAAACTGAAGAAAAATCGCCGACAGAAGAAGGC 1617

Qy       1594 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1653
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1677

Qy       1654 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCAAAAAACGGAACGGCGCAT 1713
             |||||||||||||||||||||||||||||||||||||||||  |||  ||||  || |
Db       1678 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCGCAAATTGGAAAAGCACGC 1737

Qy       1714 TATACCGGCACTTGGGAAGCGCGTATCGGCGTATCGGATA-------------------- 1753
             |||||||||||||||||||||||||||||||||||||  ||||||
Db       1738 TATACCGGCACTTGGGAAGCGCGTATCGGCGTGCCGGATAAGAAAGGCGAACAGCTAGAT 1797

Qy       1754 -GTGGTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGGGGCAAAAGCAGAATTTACC 1812
              |   |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||  |
Db       1798 GGCACTACGTCCATTCAAAAGGATAGCTATGCGAATCAAGCGGCAAAAGCAGAATTTGAC 1857

Qy       1813 GTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAATGATACAACCCCC 1872
             ||||||||  |  |||||| || | || || | | ||||||||||||||||||||  ||||
Db       1858 GTTGATTTTGGTGCGAAGTCGCTTTCAGGTAAGTTGACAGAAAAAAATGATACACACCCC 1917

Qy       1873 GCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCTTTGGCGCATACT 1932
             |||||||||||||||||||||||||||||||| || |||||||||||||||||||| ||||
Db       1918 GCTTTTTATATTGAAAAAGGTGTGATTGATGGCAACGGTTTCCACGCTTTGGCGCGTACT 1977

Qy       1933 CGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAAGCC 1992
             || || ||  ||| ||||   | |||||||| |||||||||| ||  | |  || ||||||
Db       1978 CGTGAAAATGGTGTTGATTTGTCTGGGCAAGGTTCGACTAATCCCCAAAGTTTTAAAGCC 2037

Qy       1993 GACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGTAAT 2052
                |||||||| |||  ||| || |||||||| |||||||||||| ||||| ||||||
Db       2038 AGTAATCTTCTCGTAGAAGGAGGATTTTATGGTCCGCAGGCGGCAGAGTTGGGTGGTAAT 2097

Qy       2053 ATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGACAAGGAG 2112
             |||||||||||| |||||||| |||| | |||||| || ||||| ||||||    |||||
Db       2098 ATTATCGACAGTGACCGGAAAATCGGCGTGGTATTCGGTGCGAAGAAAGATATGCAGGAG 2157

Qy       2113 GCAACACGATGA 2124
              |    |  ||||
Db       2158 GTGGAAAAATGA 2169
```

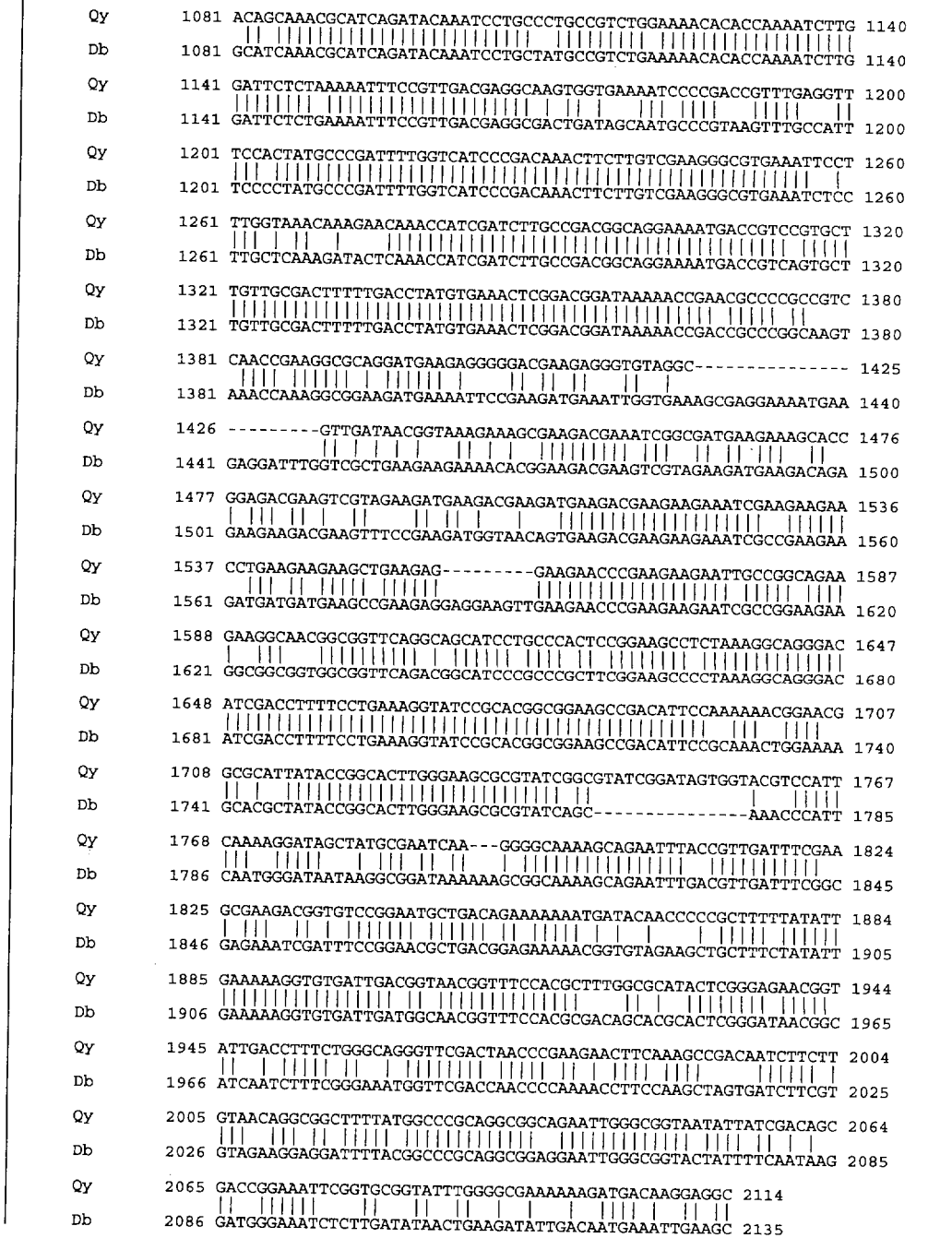
RESULT 3
ACA41945
ID ACA41945 standard; DNA; 2226 BP.

XX
AC ACA41945;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #23602.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Neisseria meningitidis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU38075.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 29815; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2226 BP; 701 A; 466 C; 566 G; 493 T; 0 U; 0 Other;

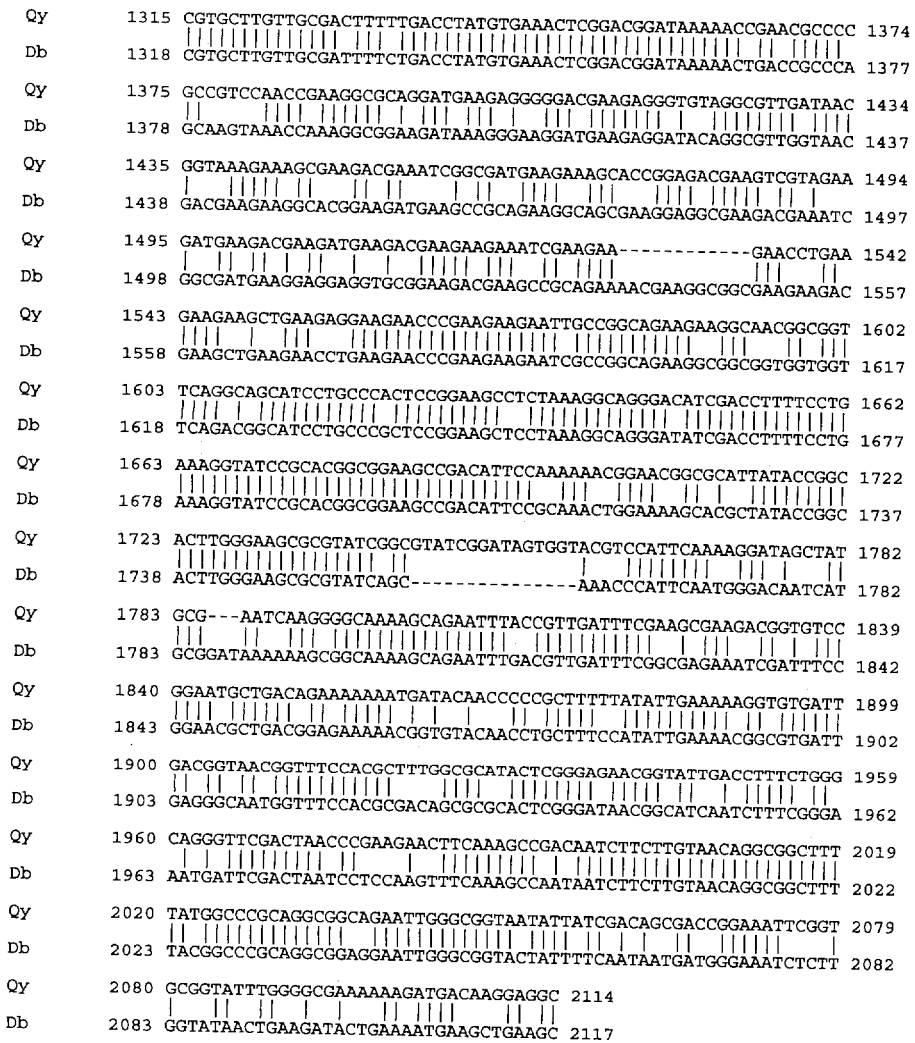
Query Match 71.4%; Score 1516.8; DB 7; Length 2226;

Best Local Similarity 83.3%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 307; Indels 51; Gaps 4;





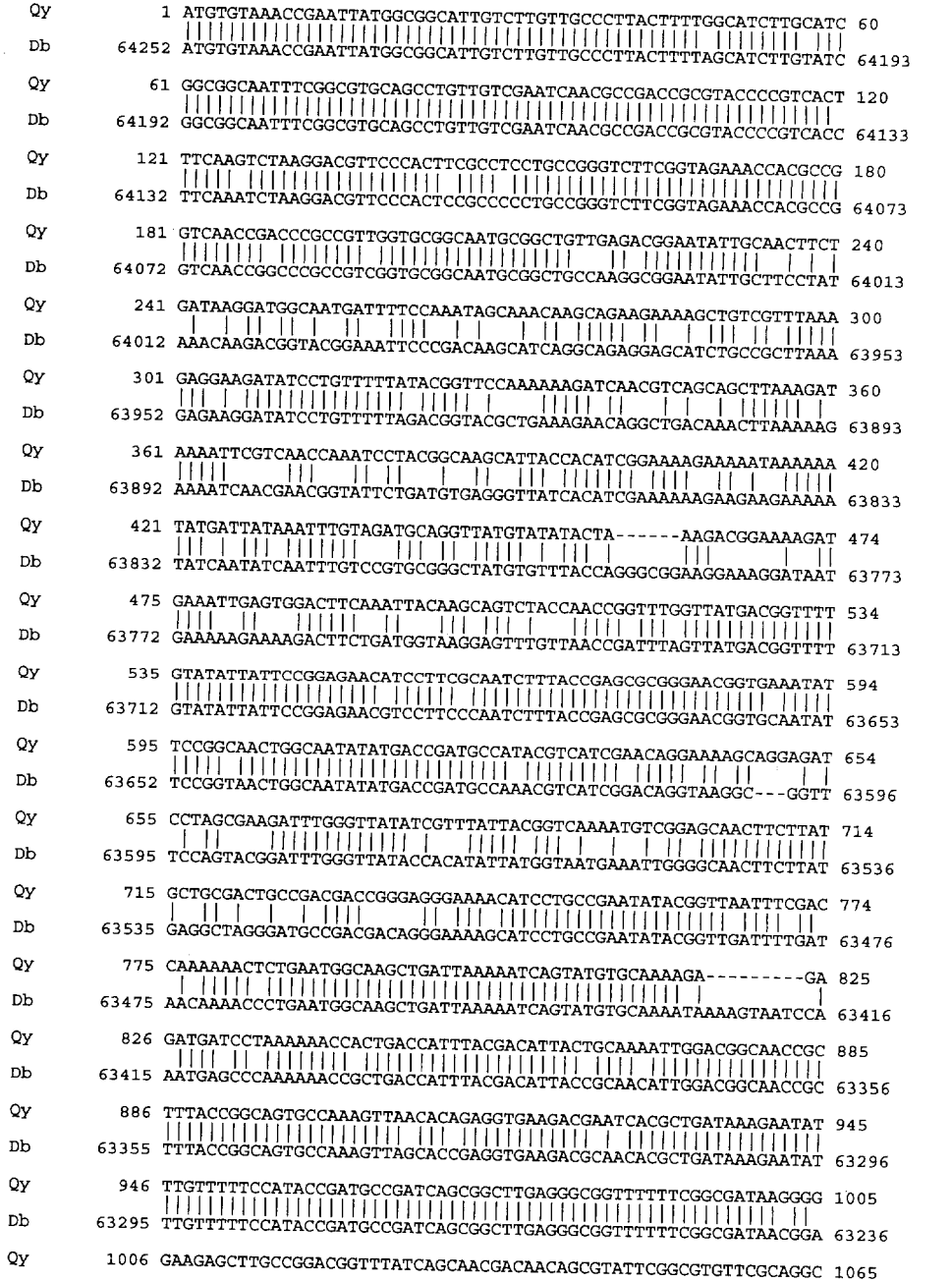
```
RESULT 4
AAX23321
ID   AAX23321 standard; cDNA; 2226 BP.
XX
AC   AAX23321;
XX
DT   11-JUN-1999  (first entry)
XX
DE   N. meningitidis strain H44/76 LbpB cDNA.
XX
KW   LbpB; lactoferrin binding protein; vaccine; neisserial disease;
KW   meningitis; diagnosis; treatment; ds.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2226
FT                   /*tag= a
FT                   /product= "LbpB"
XX
PN   WO9909176-A1.
XX
PD   25-FEB-1999.
XX
PF   10-AUG-1998;    98WO-EP005117.
XX
PR   15-AUG-1997;    97GB-00017423.
PR   05-FEB-1998;    98GB-00002544.
XX
PA   (UYUT-) RIJKSUNIV UTRECHT.
PA   (TECH-) TECHNOLOGY FOUND TECHNOLOGIESTICHTING ST.
XX
PI   Pettersson-Fernholm AM,  Tommassen JPM;
XX
DR   WPI; 1999-190165/16.
DR   P-PSDB; AAW93494.
XX
PT   New lactoferrin-binding protein B polynucleotides - obtained from
PT   Neisseria meningitidis, used to develop products for the diagnosis,
PT   prevention and treatment of neisserial disease, e.g. meningitis.
XX
PS   Claim 2; Page 90-94; 116pp; English.
XX
CC   This invention describes novel lactoferrin-binding protein B (LbpB)
CC   strains of Neisseria meningitidis. The products of this invention can be
CC   used for vaccinating humans against neisserial disease e.g. meningitis.
CC   Antibodies raised against the proteins of the invention can be used for
CC   diagnosing or treating neisserial disease in humans. The LbpB
CC   polypeptides can also be used for identifying compounds which inhibit the
CC   polypeptides
XX
SQ   Sequence 2226 BP; 683 A; 483 C; 571 G; 489 T; 0 U; 0 Other;

  Query Match          70.6%;  Score 1499.8;  DB 2;  Length 2226;
  Best Local Similarity  83.3%;  Pred. No. 0;
  Matches 1779;  Conservative    0;  Mismatches  317;  Indels   39;  Gaps    5;

Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGTATT 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy        121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
              ||||||||||||||||||||||| |||| ||||||    ||||  ||||||||||||||
Db        121 TTCAAGTCTAAGGACGTTCCCACTCCGCCCCCTGCCAAACCTTCTATAGAAACCACGCCG 180

Qy        181 GT------CAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCA 234
              ||      | | | || ||||| ||||||||||||||||||||||| |||| | | |||
Db        181 GTGCCGTCAACCGGGCCTGCCGTCGGTGCGGCAATGCGGCTGTTGAGGCGGATTTTCGCA 240

Qy        235 ACTTCTGATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCG 294
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTTCTGATAAGGTTGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCG 300

Qy        295 TTTAAAGAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTT 354
              |||||||| |  ||| | |||||||||||||||| ||||||||| ||| ||||  ||||
Db        301 TTTAAAGAAGGTGATGTTCTGTTTTTATACGGTTCAAAAAAAGATAAACTTCAGTGGCTT 360

Qy        355 AAAGATAAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAAT 414
              || |||||||||| |||||  |||||||  | |  ||||  |||||  |||||  ||||||
Db        361 AAGGATAAAATTCATCAACGCAATCCTAATGTAGAAATTAGGACATCAGAAAATGAAAAT 420

Qy        415 AAAAAATATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGAT 474
              |||||||||| |||| |||||||  |||||  ||||||||||||||||||| ||||||  ||||
Db        421 AAAAAATATGGTTATGAATTTGTGGATGCCGGTTATGTATATACTAAAAACGGAACAGAT 480

Qy        475 GAAATTGAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTT 534
              |||||||||||||||||||||||  ||||||||  | | ||  ||  |||||  ||  ||||||||||
Db        481 GAAATTGAGTGGACTTCAAATCGCAAGCAGTTTTCTAATCGTTTTGGCTACGACGGTTTT 540

Qy        535 GTATATTATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATAT 594
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||  |||||
Db        541 GTATATTATTCCGGAGAACATCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGCAATAT 600

Qy        595 TCCGGCAACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGAT 654
              |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCGGTAACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGAT 660

Qy        655 CCTAGCGAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTAT 714
              |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        661 CCTAGCGAAGATTTGGGTTATCTCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTAT 720

Qy        715 GCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGAC 774
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db        721 GCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTGATTTCGAT 780

Qy        775 CAAAAAACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCT 834
              | ||||||  |||  ||  |  | ||||||||||||||||||||||  |  |||
Db        781 AAGAAAACTTTGACGGGTCAATTAATTAAAAATCAGTATGTGCAAAAGAAAACCGATGAA 840

Qy        835 AAAAAACCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGC 894
              || |||||||||||||||||||||||||||||  |||| |||||||||||||||||||||||
Db        841 AAGAAACCACTGACCATTTACGACATTACCGCAACATTGGACGGCAACCGCTTTACCGGC 900

Qy        895 AGTGCCAAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTC 954
              |||||||||||||||| ||| |||||||||  |||||||||||||  ||||||||||||
Db        901 AGTGCCAAAGTTAACACCGAGTTGAAGACGAGCCACGCTGATAAAGAGCATTTGTTTTTC 960

Qy        955 CATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTT 1014
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTT 1020

Qy       1015 GCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAA 1074
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTATTCGCAGGCA---AAAAA 1077

Qy       1075 ACAGAGACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAA 1134
              |||  ||  |||||||| ||||||||||||||||  |||||||| ||||||||||||||
Db       1078 ACAAACGCATCAAACGCAGCAGATACAAATCCTGCTATGCCGTCTGAAAAACACACCAAA 1137

Qy       1135 ATCTTGGATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTT 1194
              |||||||||||||| |||||||||||||||||||||||| | | |||| |||| ||||||
Db       1138 ATCTTGGATTCTCTGAAAATTTCCGTTGACGAGGCGACGGATAAAAATGCCCGCCCGTTT 1197

Qy       1195 GAGGTTTCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAA 1254
              |  |||||  || |||||||||||||  |||||||||||||||||||||||||||||||||
Db       1198 GCCATTTCCCCTCTGCCCGATTTTGGCCATCCCGACAAACTCCTTGTCGAAGGGCGTGAA 1257

Qy       1255 ATTCCTTTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTC 1314
              |||||||||| | | ||||  |||||||||  ||||||||||||||||||||||||||||||
Db       1258 ATTCCTTTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCGTC 1317
```

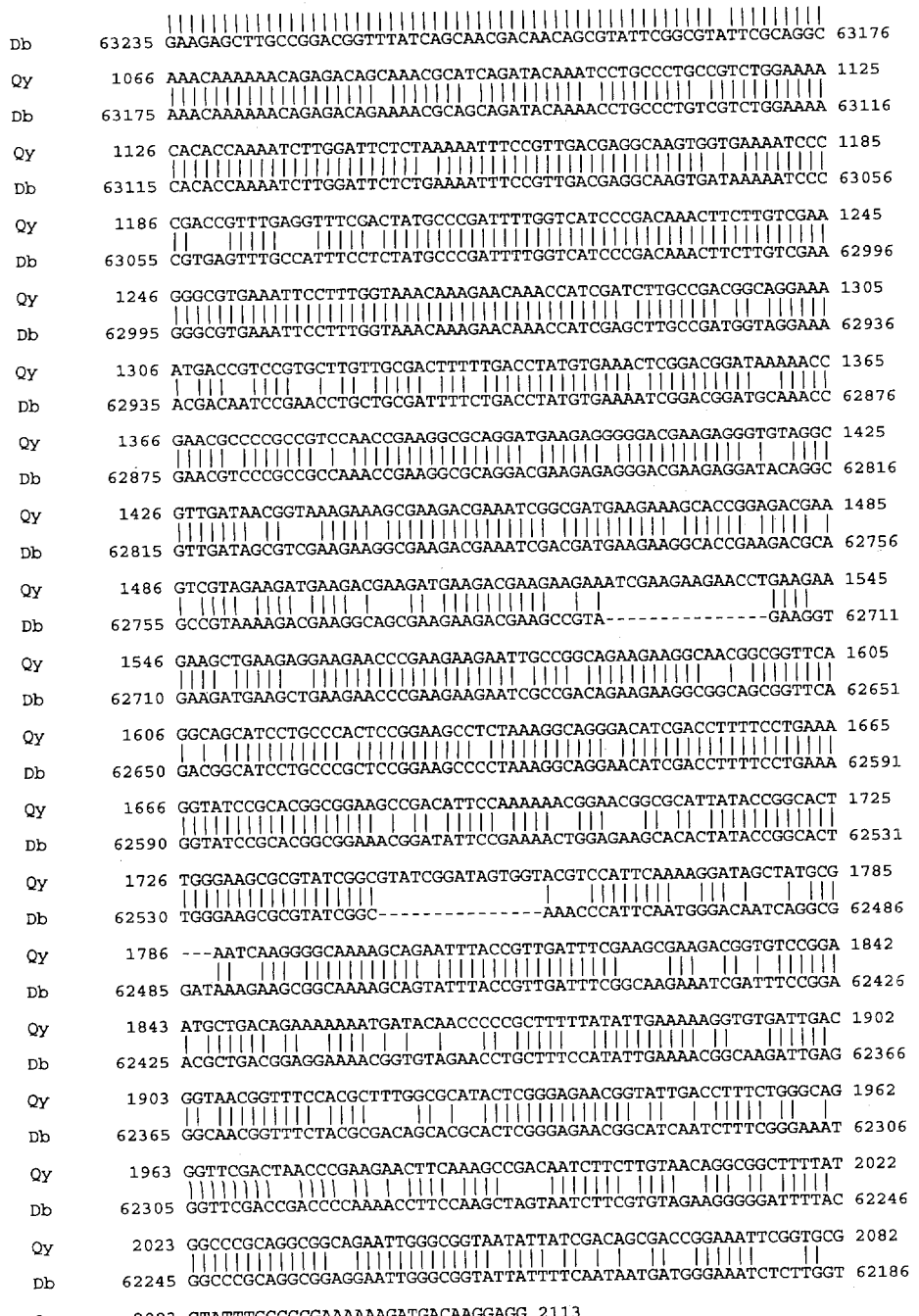


RESULT 5
AAA81489_1/c
Continuation (2 of 9) of AAA81489 from base 100001 (N. meningitidis partial DNA sequence
WP Sequence split into 9 fragments LOCUS AAA81489 Accession Aaa81489

WP	Fragment Name	Begin	End
WP	AAA81489_0	1	110000
WP	AAA81489_1	100001	210000
WP	AAA81489_2	200001	310000
WP	AAA81489_3	300001	410000
WP	AAA81489_4	400001	510000
WP	AAA81489_5	500001	610000
WP	AAA81489_6	600001	710000
WP	AAA81489_7	700001	810000
WP	AAA81489_8	800001	837096

Query Match 64.8%; Score 1375.4; DB 3; Length 110000;
Best Local Similarity 80.2%; Pred. No. 0;
Matches 1709; Conservative 0; Mismatches 371; Indels 51; Gaps 6;

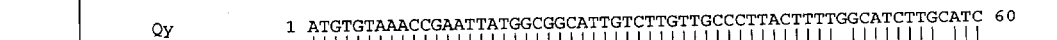


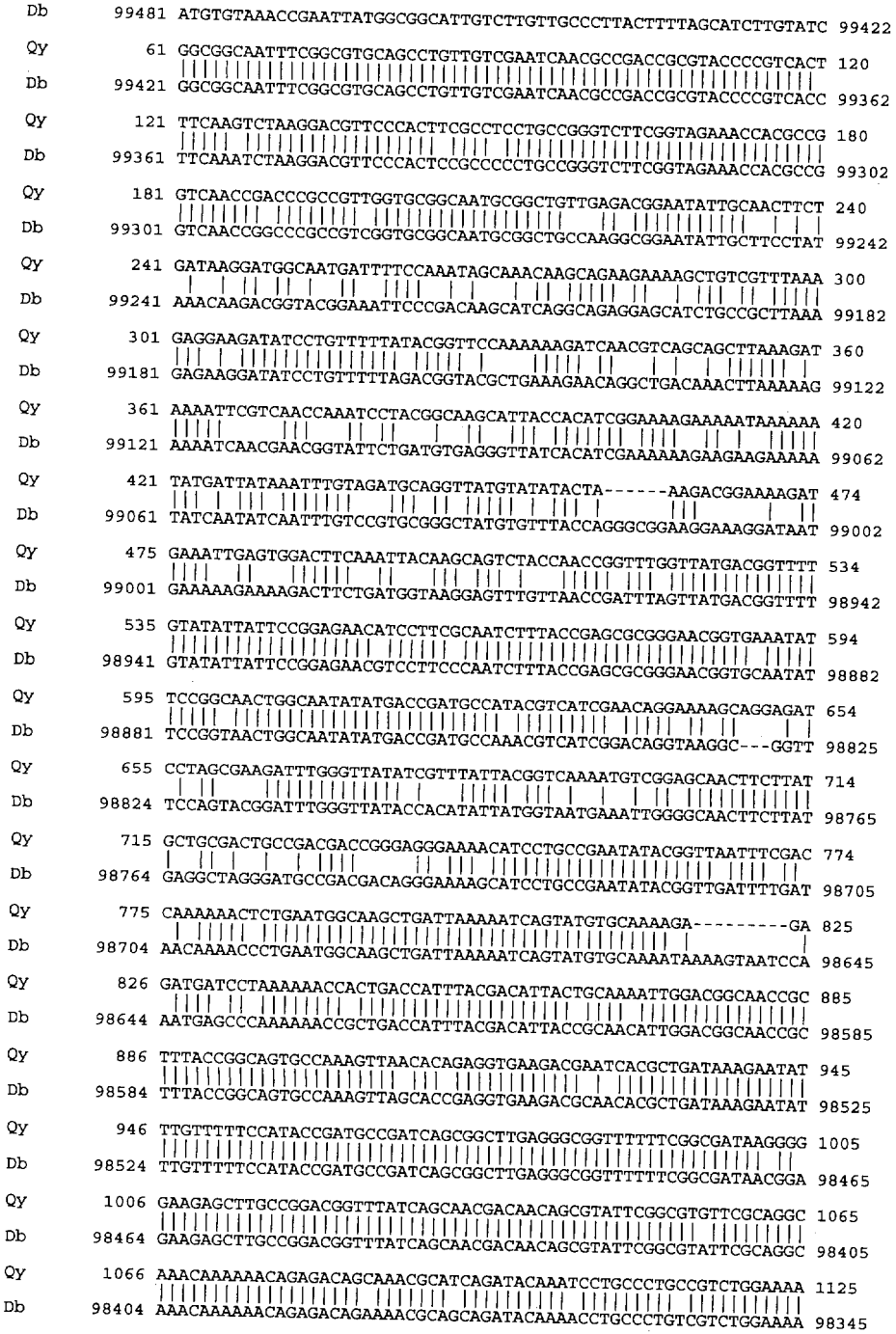


Db 62185 ATAACTGAAGGTACTGAAAATAAAGTTGATG 62155

RESULT 6
AAF21611/c
ID AAF21611 standard; DNA; 349980 BP.
XX
AC AAF21611;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US005928.
XX
PR 30-APR-1999; 99US-0132068P.
PR 08-OCT-1999; 99WO-US023573.
PR 28-FEB-2000; 2000GB-00004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections.
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of Neisseria
CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
CC represent fragments of the NMB genomic sequence, as the sequence was too
CC long to go in a record on its own it was split into 8 sequences which
CC overlap each other at the beginning and end of each sequence by 49980 bp
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
CC primers which are used in the exemplification of the present invention.
CC The NMB genome and fragments from it have antibacterial activity, and can
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
CC and/or antibodies which binds to the proteins can be used in compositions
CC for treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find antigenic
CC or immunogenic proteins which are more effective in vaccines than the
CC outer membrane proteins currently used
XX
SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;

Query Match 64.8%; Score 1375.4; DB 3; Length 349980;
Best Local Similarity 80.2%; Pred. No. 0;
Matches 1709; Conservative 0; Mismatches 371; Indels 51; Gaps 6;



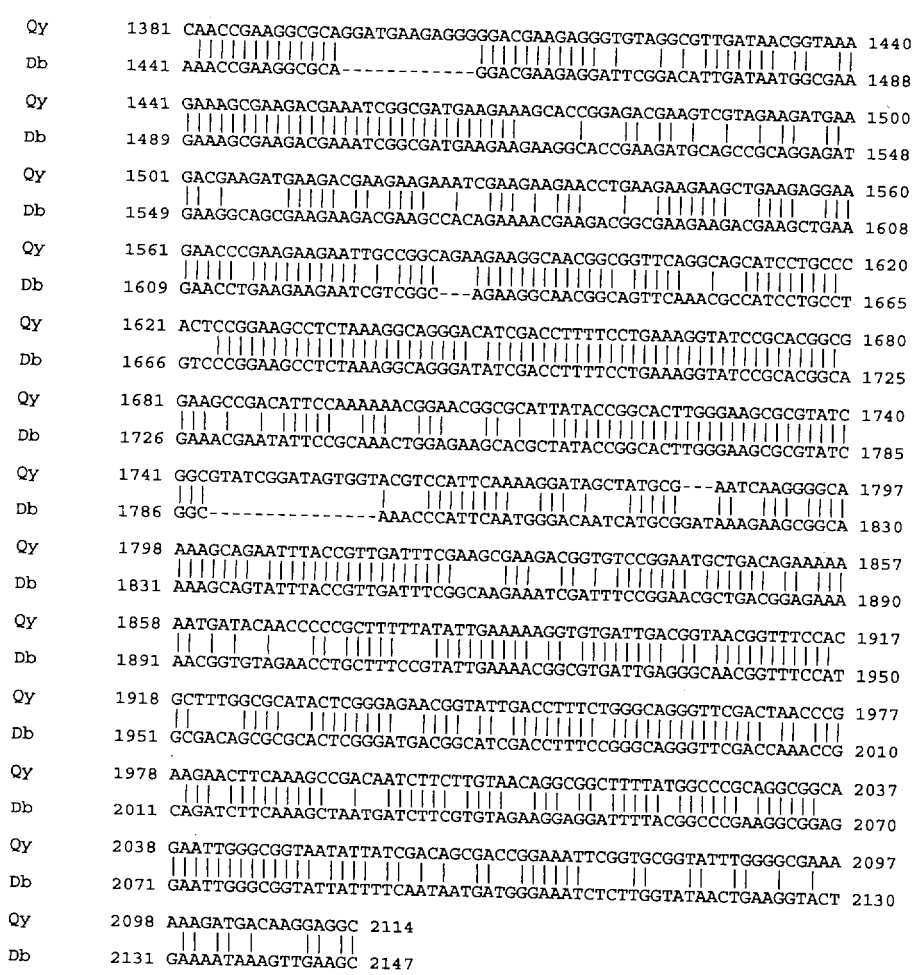


RESULT 7
AAX23319
ID AAX23319 standard; cDNA; 2277 BP.

XX
AC AAX23319;
XX
DT 11-JUN-1999 (first entry)
XX
DE N. meningitidis strain BNCV LbpB cDNA.
XX
KW LbpB; lactoferrin binding protein; vaccine; neisserial disease;
KW meningitis; diagnosis; treatment; ds.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 100. .2277
FT /*tag= a
FT /product= "LbpB"
XX
PN WO9909176-A1.
XX
PD 25-FEB-1999.
XX
PF 10-AUG-1998; 98WO-EP005117.
XX
PR 15-AUG-1997; 97GB-00017423.
PR 05-FEB-1998; 98GB-00002544.
XX
PA (UYUT-) RIJKSUNIV UTRECHT.
PA (TECH-) TECHNOLOGY FOUND TECHNOLOGIESTICHTING ST.
XX
PI Pettersson-Fernholm AM, Tommassen JPM;
XX
DR WPI; 1999-190165/16.
DR P-PSDB; AAW93492.
XX
PT New lactoferrin-binding protein B polynucleotides - obtained from
PT Neisseria meningitidis, used to develop products for the diagnosis,
PT prevention and treatment of neisserial disease, e.g. meningitis.
XX
PS Claim 2; Page 74-78; 116pp; English.
XX
CC This invention describes novel lactoferrin-binding protein B (LbpB)
CC strains of Neisseria meningitidis. The products of this invention can be
CC used for vaccinating humans against neisserial disease e.g. meningitis.
CC Antibodies raised against the proteins of the invention can be used for
CC diagnosing or treating neisserial disease in humans. The LbpB
CC polypeptides can also be used for identifying compounds which inhibit the
CC polypeptides
XX
SQ Sequence 2277 BP; 700 A; 469 C; 573 G; 535 T; 0 U; 0 Other;

Query Match 61.0%; Score 1296.2; DB 2; Length 2277;
Best Local Similarity 79.0%; Pred. No. 5.7e-312;
Matches 1672; Conservative 0; Mismatches 373; Indels 72; Gaps 8;

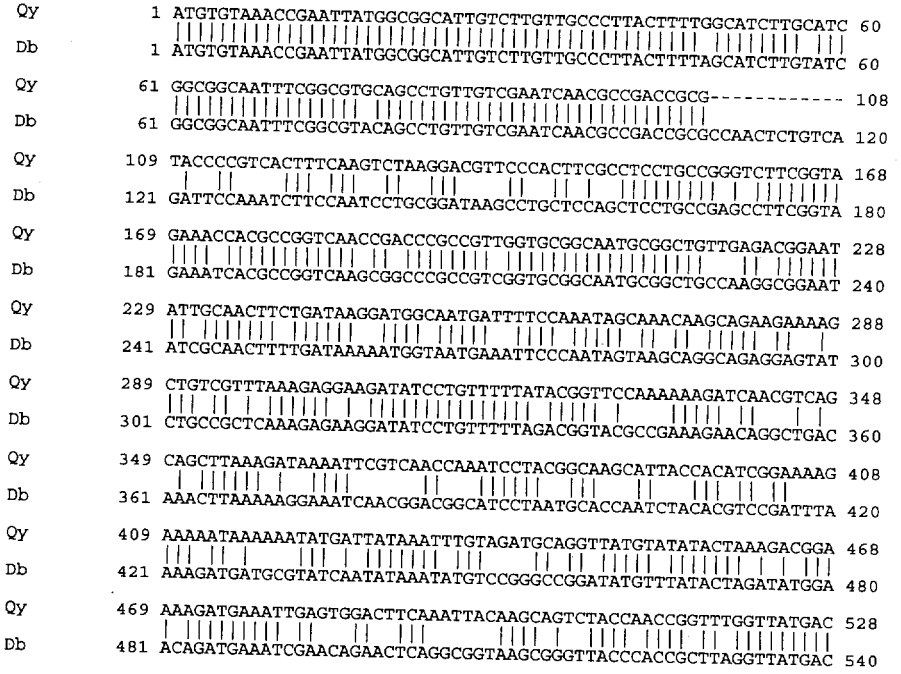
```
Qy        1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db      100 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGTATC 159
Qy       61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db      160 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 219
Qy      121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db      220 TTCAAGTCTAAGGACGTTCCCACTCCGCCCCCTGCCAAACCTTCTATAGAAATCACGCCG 279
Qy      181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
Db      280 GTCAACCGGCCCGCCGTCGGTGCGGCAATGCGGCTGCCAAGGCGGAATACTGCTTTTCAT 339
Qy      241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
Db      340 CGTGAAGATGGCACGGAAATTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTCAA 399
Qy      301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
Db      400 GAAGGTGATGTTCTGTTTTTATACGGTTCAAAAGGAAATAAACTTCAACAACTTAAAAGC 459
Qy      361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
Db      460 GAAATTCATAAACGTGATTCCGATGTAGAAATTAGGACATCAGAAAAGGAAAATAAAAAA 519
Qy      421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
Db      520 TATGATTATAAATTTGTAGATGCAGGTTATGTATA---TGTAAAGGGAAAAGATGAAATT 576
Qy      481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
Db      577 AAGTGGACTTCAGATTACAAGCAGTTTTCCAACCGCTTAGGTTATGACGGTTTTGTATAT 636
Qy      541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
Db      637 TATTCCGGAGAACGTCCTTCCCAATCTTTACCGAGTGCGGGAACGGTGGAATATTCTGGT 696
Qy      601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
Db      697 AACTGGCAATATATGACCGATGCCAAACGTCATCGAGCAGGTAAGGC---GGTTGGCATT 753
Qy      661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
Db      754 GACAATTTGGGTTATTACACATTTTATGGTAACGATGTTGGTGCAACTTCTTATGCGGCT 813
Qy      721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
Db      814 AAGGATGTCGACGAAAGGGAAAAACATCCTGCTAAATATACGGTAGATTTCGGTAACAAA 873
Qy      781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
Db      874 ACCCTGACGGGCGAGCTGATTAAAAACCAATATGTCAAACCCAGTGAGAAGC---AAAAA 930
Qy      841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
Db      931 CCGCTGACCATTTACAACATCACTGCCGATTTAAACGGCAACCGCTTTACCGGCAGTGCC 990
Qy      901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
Db      991 AAGGTCAATCCTGATTTAGCGAAAAGCCATGCCAATAAGGAGCATTTGTTTTTCCATGCC 1050
Qy      961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
Db     1051 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1110
Qy     1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
Db     1111 CGGTTTATCAGCAACGACAACAGCGTATTCGGTGTATTCGCAGGCAAACAAAATAGCC-- 1168
Qy     1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
Db     1169 ----------------------------CCGTGCCGTCTGGAAAACACACCAAAATCTTG 1200
Qy     1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
Db     1201 GATTCTCTGAAAATTTCCGTTGATGAGGCAAGTGGTGAAAATCCCCGACCGTTTGCCATT 1260
Qy     1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
Db     1261 TCTCCTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCATGAAATTCCT 1320
Qy     1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
Db     1321 TTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCGTCAGTGCT 1380
Qy     1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
Db     1381 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGCC 1440
```

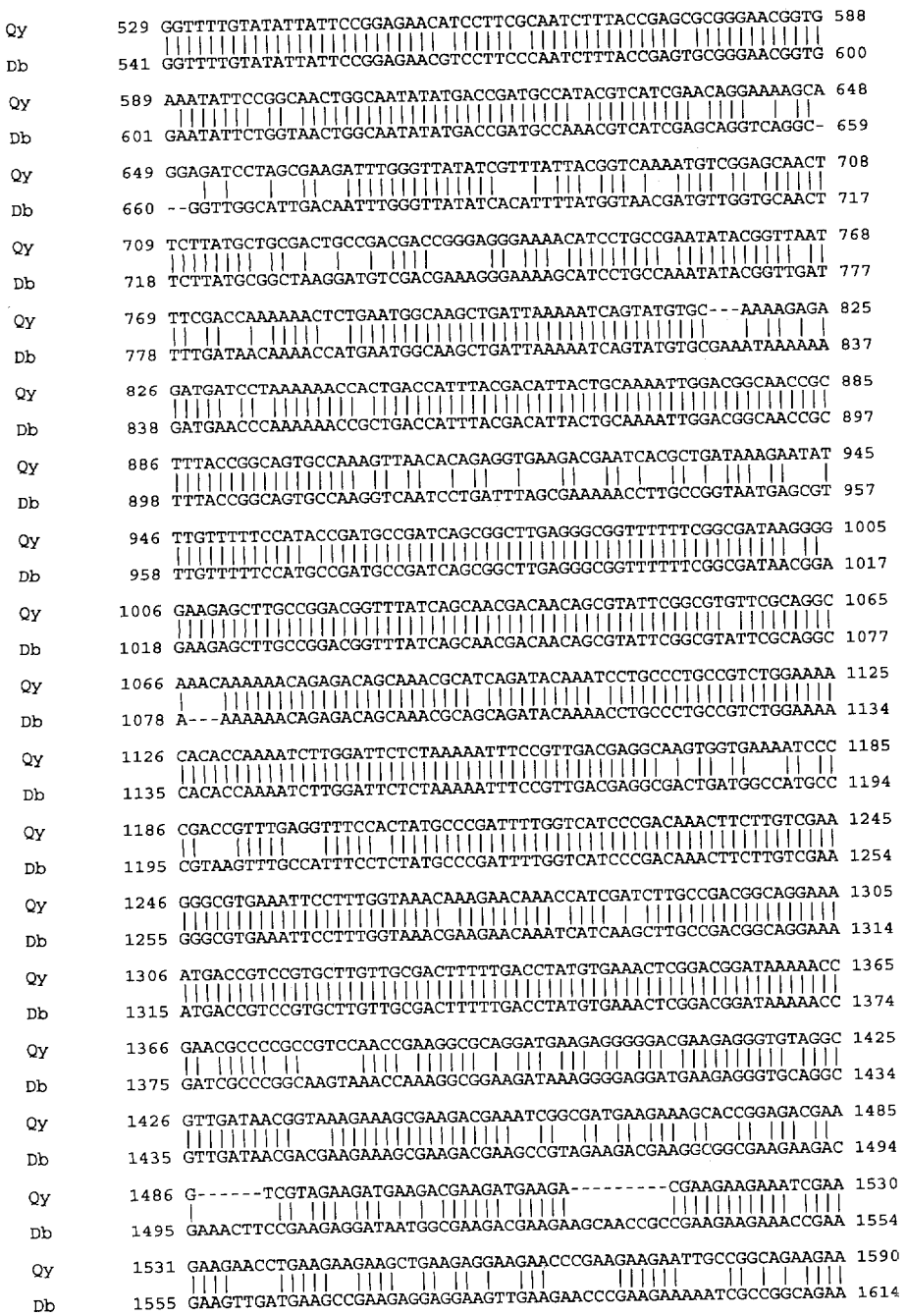


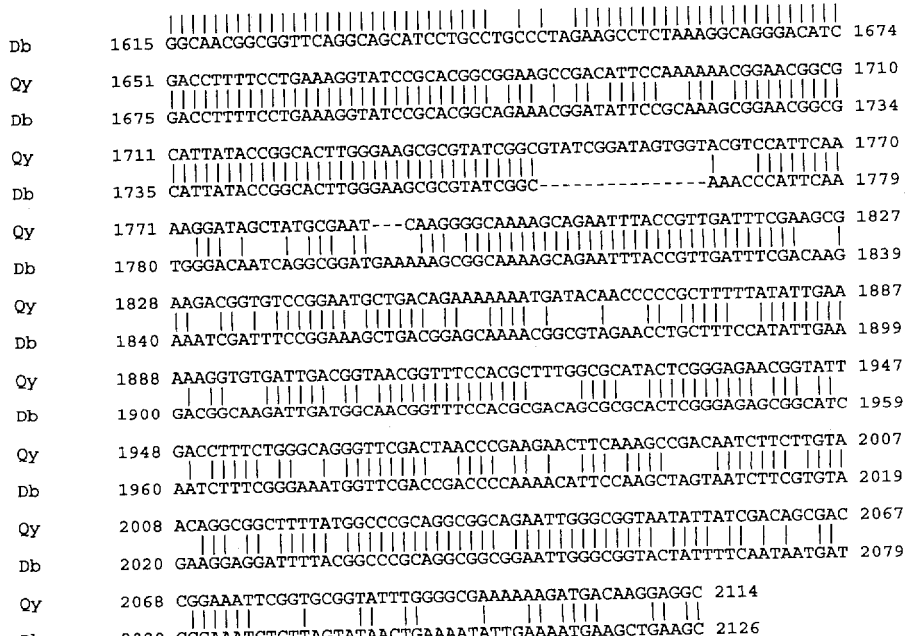
RESULT 8
AAX23322
ID AAX23322 standard; cDNA; 2262 BP.
XX
AC AAX23322;
XX
DT 11-JUN-1999 (first entry)
XX
DE N. meningitidis strain M990 LbpB cDNA.
XX
KW LbpB; lactoferrin binding protein; vaccine; neisserial disease;
KW meningitis; diagnosis; treatment; ds.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1. .2262
FT /*tag= a
FT /product= "LbpB"
XX

PN WO9909176-A1.
XX
PD 25-FEB-1999.
XX
PF 10-AUG-1998; 98WO-EP005117.
XX
PR 15-AUG-1997; 97GB-00017423.
PR 05-FEB-1998; 98GB-00002544.
XX
PA (UYUT-) RIJKSUNIV UTRECHT.
PA (TECH-) TECHNOLOGY FOUND TECHNOLOGIESTICHTING ST.
XX
PI Pettersson-Fernholm AM, Tommassen JPM;
XX
DR WPI; 1999-190165/16.
DR P-PSDB; AAW93495.
XX
PT New lactoferrin-binding protein B polynucleotides - obtained from
PT Neisseria meningitidis, used to develop products for the diagnosis,
PT prevention and treatment of neisserial disease, e.g. meningitis.
XX
PS Claim 2; Page 98-102; 116pp; English.
XX
CC This invention describes novel lactoferrin-binding protein B (LbpB)
CC strains of Neisseria meningitidis. The products of this invention can be
CC used for vaccinating humans against neisserial disease e.g. meningitis.
CC Antibodies raised against the proteins of the invention can be used for
CC diagnosing or treating neisserial disease in humans. The LbpB
CC polypeptides can also be used for identifying compounds which inhibit the
CC polypeptides
XX
SQ Sequence 2262 BP; 698 A; 487 C; 590 G; 487 T; 0 U; 0 Other;

Query Match 60.8%; Score 1291.8; DB 2; Length 2262;
Best Local Similarity 78.1%; Pred. No. 7.1e-311;
Matches 1676; Conservative 0; Mismatches 417; Indels 54; Gaps 8;







RESULT 9
AAA81482/c
ID AAA81482 standard; DNA; 14652 BP.
XX
AC AAA81482;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_30 SEQ ID NO:30.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-0103794P.
PR 30-APR-1999; 99US-0132068P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX

PS Claim 7; Page 582-586; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neissariae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 14652 BP; 3745 A; 4117 C; 3436 G; 3354 T; 0 U; 0 Other;

Query Match 44.1%; Score 937.4; DB 3; Length 14652;
Best Local Similarity 81.2%; Pred. No. 2.3e-222;
Matches 1155; Conservative 0; Mismatches 226; Indels 42; Gaps 4;

```
Qy      703 GCAACTTCTTATGCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACG 762
Db    14652 GCAACTTCTTATGAGGCTAGGGATGCCGACGACAGGGAAAAGCATCCTGCCGAATATACG 14593

Qy      763 GTTAATTTCGACCAAAAAACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAG 822
Db    14592 GTTGATTTTGATAACAAAACCCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAT 14533

Qy      823 A---------GAGATGATCCTAAAAAACCACTGACCATTTACGACATTACTGCAAAATTG 873
Db    14532 AAAAGTAATCCAAATGAGCCCAAAAAACCGCTGACCATTTACGACATTACCGCAACATTG 14473

Qy      874 GACGGCAACCGCTTTACCGGCAGTGCCAAAGTTAACACAGAGGTGAAGACGAATCACGCT 933
Db    14472 GACGGCAACCGCTTTACCGGCAGTGCCAAAGTTAGCACCGAGGTGAAGACGCAACACGCT 14413

Qy      934 GATAAAGAATATTTGTTTTTCCATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTC 993
Db    14412 GATAAAGAATATTTGTTTTTCCATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTC 14353

Qy      994 GGCGATAAGGGGGAAGAGCTTGCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGC 1053
Db    14352 GGCGATAACGGAGAAGAGCTTGCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGC 14293

Qy     1054 GTGTTCGCAGGCAAACAAAAAACAGAGACAGCAAACGCATCAGATACAAATCCTGCCCTG 1113
Db    14292 GTATTCGCAGGCAAACAAAAAACAGAGACAGAAAACGCAGCAGATACAAAACCTGCCCTG 14233

Qy     1114 CCGTCTGGAAAACACACCAAAATCTTGGATTCTCTAAAAATTTCCGTTGACGAGGCAAGT 1173
Db    14232 TCGTCTGGAAAACACACCAAAATCTTGGATTCTCTGAAAATTTCCGTTGACGAGGCAAGT 14173

Qy     1174 GGTGAAAATCCCCGACCGTTTGAGGTTTCCACTATGCCCGATTTTGGTCATCCCGACAAA 1233
Db    14172 GATAAAAATCCCCGTGAGTTTGCCATTTCCTCTATGCCCGATTTTGGTCATCCCGACAAA 14113

Qy     1234 CTTCTTGTCGAAGGGCGTGAAATTCCTTTGGTAAACAAAGAACAAACCATCGATCTTGCC 1293
Db    14112 CTTCTTGTCGAAGGGCGTGAAATTCCTTTGGTAAACAAAGAACAAACCATCGAGCTTGCC 14053

Qy     1294 GACGGCAGGAAAATGACCGTCCGTGCTTGTTGCGACTTTTTGACCTATGTGAAACTCGGA 1353
Db    14052 GATGGTAGGAAAACGACAATCCGAACCTGCTGCGATTTTCTGACCTATGTGAAAATCGGA 13993

Qy     1354 CGGATAAAAACCGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAA 1413
Db    13992 CGGATGCAAACCGAACGTCCCGCCGCCAAACCGAAGGCGCAGGACGAAGAGAGGGACGAA 13933

Qy     1414 GAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGC 1473
Db    13932 GAGGATACAGGCGTTGATAGCGTCGAAGAAGGCGAAGACGAAATCGACGATGAAGAAGGC 13873

Qy     1474 ACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAA 1533
Db    13872 ACCGAAGACGCAGCCGTAAAAGACGAAGGCAGCGAAGAAGACGAAGCCGTA--------- 13822

Qy     1534 GAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGC 1593
Db    13821 ------GAAGGTGAAGATGAAGCTGAAGAACCCGAAGAAGAATCGCCGACAGAAGAAGGC 13768

Qy     1594 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1653
Db    13767 GGCAGCGGTTCAGACGGCATCCTGCCCGCTCCGGAAGCCCCTAAAGGCAGGAACATCGAC 13708

Qy     1654 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCAAAAAACGGAACGGCGCAT 1713
Db    13707 CTTTTCCTGAAAGGTATCCGCACGGCGGAAACGGATATTCCGAAAACTGGAGAAGCACAC 13648

Qy     1714 TATACCGGCACTTGGGAAGCGCGTATCGGCGTATCGGATAGTGGTACGTCCATTCAAAAG 1773
Db    13647 TATACCGGCACTTGGGAAGCGCGTATCGGC--------------AAACCCATTCAATGG 13603

Qy     1774 GATAGCTATGCG---AATCAAGGGGCAAAAGCAGAATTTACCGTTGATTTCGAAGCGAAG 1830
Db    13602 GACAATCAGGCGGATAAAGAAGCGGCAAAAGCAGTATTTACCGTTGATTTCGGCAAGAAA 13543

Qy     1831 ACGGTGTCCGGAATGCTGACAGAAAAAAATGATACAACCCCCGCTTTTTATATTGAAAAA 1890
Db    13542 TCGATTTCCGGAACGCTGACGGAGGAAAACGGTGTAGAACCTGCTTTCCATATTGAAAAC 13483

Qy     1891 GGTGTGATTGACGGTAACGGTTTCCACGCTTTGGCGCATACTCGGGAGAACGGTATTGAC 1950
Db    13482 GGCAAGATTGAGGGCAACGGTTTCTACGCGACAGCACGCACTCGGGAGAACGGCATCAAT 13423

Qy     1951 CTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAAGCCGACAATCTTCTTGTAACA 2010
Db    13422 CTTTCGGGAAATGGTTCGACCGACCCCAAAACCTTCCAAGCTAGTAATCTTCGTGTAGAA 13363

Qy     2011 GGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGTAATATTATCGACAGCGACCGG 2070
Db    13362 GGGGGATTTTACGGCCCGCAGGCGGAGGAATTGGGCGGTATTATTTTCAATAATGATGGG 13303

Qy     2071 AAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGACAAGGAGG 2113
Db    13302 AAATCTCTTGGTATAACTGAAGGTACTGAAAATAAAGTTGATG 13260
```

RESULT 10
AAF91389
ID AAF91389 standard; DNA; 1000 BP.
XX
AC AAF91389;
XX
DT 04-MAY-2001 (first entry)
XX
DE N. meningitidis (serogroup B) LbpA gene upstream sequence, SEQ ID:15.
XX
KW Modified Gram-negative bacterium; outer membrane vesicle; bleb; vaccine;
KW genetically modified; protective antigen expression; LPS detoxification;
KW LPS; lipid A; homologous recombination vector; immunisation;
KW immunoprotective; non-toxic; paediatric; cyclic; circular; ds.
XX

OS Neisseria meningitidis.
XX
PN WO200109350-A2.
XX
PD 08-FEB-2001.
XX
PF 31-JUL-2000; 2000WO-EP007424.
XX
PR 03-AUG-1999; 99GB-00018319.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Berthet FJ, Dalemans WLJ, Denoel P, Dequesne G, Feron C, Lobet Y;
PI Poolman J, Thiry G, Thonnard J, Voet P;
XX
DR WPI; 2001-138654/14.
XX
PT New isolated polynucleotide useful for outer membrane vesicle preparation
PT from Gram-negative bacterial strain for vaccination of microbial
PT infections.
XX
PS Claim 46; Page 81; 128pp; English.
XX
CC The invention relates to a genetically-engineered outer membrane vesicle
CC (bleb) preparation from a Gram-negative bacterium for use as a vaccine.
CC The blebs of the invention are improved with respect to their
CC immunogenicity and toxicity by the introduction of one or more genetic
CC changes to the chromosome of the bacterium from which the blebs are
CC derived. The changes made include the upregulation of protective antigen
CC expression, the downregulation of immunodominant non-protective antigen
CC expression, and genetic changes which result in detoxification of the
CC Lipid A moiety of lipopolysaccharide (LPS). The invention also
CC encompasses modified Gram-negative bacterial strains from which the bleb
CC preparations are made, a vector suitable for performing recombination
CC events (for the generation of the modified bacterial strains),
CC bacterially-derived nucleic acid sequences used in such a vector, and an
CC immunoprotective and non-toxic Gram-negative bleb, ghost, or killed whole
CC cell vaccine suitable for paediatric use. The bleb preparation is useful
CC in the manufacture of a medicament for immunising a human host against a
CC disease caused by infection of one or more of the following: Neisseria
CC meningitidis, Neisseria gonorrhoeae, Haemophilus influenza, Moraxella
CC catarrhalis, Pseudomonas aeruginosa, Chlamydia trachomatis, and Chlamydia
CC pneumonia. The invention may also be used to provide immunisation against
CC the influenza virus. Bacterially derived nucleotide sequences of the
CC invention are used in the performance of homologous recombination events
CC up to 1000 bp upstream of a bacterial chromosomal gene in order to either
CC increase or decrease expression of that gene. Immunoprotective and non-
CC toxic Gram-negative bleb, ghost, or killed whole cell vaccines are more
CC immunogenic, less toxic and safer, and are particularly useful for
CC paediatric use. The present sequence represents a specifically claimed
CC Neisseria meningitidis nucleic acid sequence
XX
SQ Sequence 1000 BP; 328 A; 203 C; 278 G; 191 T; 0 U; 0 Other;

Query Match 23.4%; Score 497.6; DB 4; Length 1000;
Best Local Similarity 75.8%; Pred. No. 2e-113;
Matches 679; Conservative 0; Mismatches 189; Indels 28; Gaps 4;

```
Qy       1213 GATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCTTTGGTAAACAAA 1272
             ||||||||||||||||||| |||||||||||||||||||||||||||||| | | ||
Db          1 GATTTTGGTCATCCCGACAAGCTTCTTGTCGAAGGGCGTGAAATTCCTTTGGTTAGCCAA 60

Qy       1273 GAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCTTGTTGCGACTTT 1332
             ||  |||||||| | ||||||||| |||||| ||||||||||||||||||||||||||||
Db         61 GAGAAAACCATCAAGCTTGCCGATGGCAGGGAAATGACCGTCCGTGCTTGTTGCGACTTT 120

Qy       1333 TTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTCCAACCGAAGGCG 1392
             |||||||||||||||||||||||||||||||||||||||||||| ||      |||| ||||||
Db        121 TTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCGGCAAGTAAACCAAAGGCG 180

Qy       1393 CAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGAC 1452
              | ||| ||  || ||| |||||| ||| |||||||| |||||   ||||| |||||| |
Db        181 GAAGATAAAAGGGAGGATGAAGAGAGTGCAGGCGTTGGTAACGTCGAAGAAGGCGAAGGC 240

Qy       1453 GAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAA 1512
              ||| |   ||| || ||          || |||||  | |||||    | ||||||| |||
Db        241 GAAGTTTCCGAAGATGA---------AGGCGAAGAAGCCGAAGAAATCGTCGAAGAAGAA 291

Qy       1513 GACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAA 1572
                |||||||||   ||||| |||   |  |||    ||||  ||||||  ||||||||
Db        292 CCCGAAGAAGAAGCTGAAGAGGAAGAAGCTGAACCCAAAGAAGTTGAAGAAACCGAAGAA 351

Qy       1573 GAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCC 1632
              ||| |||| |||||||| ||  | ||||||||  |  |||||||||  |  ||||||||
Db        352 AAATCGCCGACAGAAGAAAGCGGCAGCGGTTCAAACGCCATCCTGCCTGCCTCGGAAGCC 411

Qy       1633 TCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATT 1692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATT 471

Qy       1693 CCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATCGGCGTATCGGAT 1752
              |||| || |||||  || || |||||||||||||||||||||||||||||||||
Db        472 CCAAGAACCGGAAAAGCACACTATACCGGCACTTGGGAAGCGCGTATCGGC--------- 522

Qy       1753 AGTGGTACGTCCATTCAAAAGGATAGCTATGCG---AATCAAGGGGCAAAAGCAGAATTT 1809
                          ||  |||||||| ||| |   | |||   ||  ||| ||||||||||||||||
Db        523 ------ACACCCATTCAATGGGACAATCAGGCGGATAAAGAAGCGGCAAAAGCAGAATTT 576

Qy       1810 ACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAATGATACAACC 1869
              |||||| ||||||  | |||  || | |||||||  |||||| || ||||| | |  |
Db        577 ACCGTTAATTTCGGCGAGAAATCGATTTCCGGAACGCTGACGGAGAAAAACGGTGTACAA 636

Qy       1870 CCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCTTTGGCGCAT 1929
              || ||||| |||||||||||| ||  |||||| || ||||||||||||||    || |
Db        637 CCTGCTTTCTATATTGAAAACGGCAAGATTGAGGGCAACGGTTTCCACGCAACAGCACGC 696

Qy       1930 ACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAA 1989
              ||||| |||||||| ||  | ||||| ||  | |||||||| |||||  |  | |||| ||
Db        697 ACTCGTGAGAACGGCATCAATCTTTCGGGAAATGGTTCGACCAACCCCAGAACCTTCCAA 756

Qy       1990 GCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGT 2049
              ||      |||||| |||| ||| || ||||| |||||||| |||| ||||||||||||
Db        757 GCTAGTGATCTTCGTGTAGAAGGAGGATTTTACGGCCCGCA-GCGGAGGAATTGGGCGGT 815

Qy       2050 AATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGA 2105
              | |||| || | |  ||  ||||||    ||   ||  ||  |  |   || || |
Db        816 ATTATTTTCAATAAGGATGGGAAATCTCTTGGTATAACTGAAGGTACTGAAAATAA 871
```

RESULT 11
ABK37769
ID ABK37769 standard; DNA; 1000 BP.
XX
AC ABK37769;
XX
DT 08-MAY-2002 (first entry)
XX
DE DNA sequence upstream of LbpA #1 gene.
XX
KW Upstream sequence; ds; Antibacterial; vaccine; bleb;
KW Gram-negative bacteria; outer membrane; LPS; lipopolysaccharide;
KW meningitis; bacteraemia; otitis media; pneumonia; chronic bronchitis;
KW sinusitis.
XX
OS Neisseria meningitidis serogroup B.
XX
PN WO200209746-A2.
XX
PD 07-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-EP008857.
XX
PR 31-JUL-2000; 2000WO-EP007424.
PR 08-FEB-2001; 2001GB-00003170.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX
PI Berthet FJ, Dalemans W, Denoel P, Dequesne G, Feron C, Garcon N;
PI Lobet Y, Poolman J, Thiry G, Thonnard J, Voet P;
XX
DR WPI; 2002-188688/24.
XX
PT New immunogenic composition comprising an antigen derived from a pathogen
PT and a blep preparation from Neisserria meningitidis, useful as a vaccine
PT for treating or preventing disease caused by the pathogen.
XX
PS Disclosure; Page 84; 125pp; English.
XX
CC The invention relates to an immunogenic composition comprising an antigen
CC derived from a pathogen capable of protecting a host against the
CC pathogen, mixed with an adjuvant comprising a bleb preparation derived
CC from a Gram-negative bacterial strain. The immunogenic composition
CC consists of N. meningitidis B blebs or N. meningitidis C polysaccharide
CC antigen. The blebs (derived from the outer membrane) may also have their
CC toxic lipopolysaccharide (LPS) content reduced using heterologous down
CC regulating sequences for LPS pathway genes or by up regulating genes
CC involved in LPS synthesis suppression, by a promoter replacement
CC technique. The immunogenic preparation is useful in the manufacture of a
CC medicament for the treatment of a disease caused by the pathogen from
CC which the antigen is derived (e.g. from Neisseria, meningitis and
CC bacteraemia, from Moraxella, otitis media and pneumonia, and from H.
CC influenzae chronic bronchitis, sinusitis, pneumonia and otitis media).
CC The bleb derived from M. catarrhalis or from a non-typeable H. influenzae
CC is useful as an adjuvant in an immunogenic composition comprising one or
CC more pneumococcal capsular polysaccharides or protein antigens. The
CC present sequence is an upstream sequence from an N. meningitidis,
CC H.influenzae or M. catarrhalis gene involved in LPS biosynthesis, which
CC either up regulates or down regulates sequences to which it is attached
XX
SQ Sequence 1000 BP; 328 A; 203 C; 278 G; 191 T; 0 U; 0 Other;

Query Match 23.4%; Score 497.6; DB 6; Length 1000;
Best Local Similarity 75.8%; Pred. No. 2e-113;
Matches 679; Conservative 0; Mismatches 189; Indels 28; Gaps 4;

```
Qy       1213 GATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCTTTGGTAAACAAA 1272
             ||||||||||||||||||||| ||||||||||||||||||||||||||||||| | | ||
Db          1 GATTTTGGTCATCCCGACAAGCTTCTTGTCGAAGGGCGTGAAATTCCTTTGGTTAGCCAA 60

Qy       1273 GAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCTTGTTGCGACTTT 1332
             ||  |||||||| | |||||||| |||||| ||||||||||||||||||||||||||||||
Db         61 GAGAAAACCATCAAGCTTGCCGATGGCAGGGAAATGACCGTCCGTGCTTGTTGCGACTTT 120

Qy       1333 TTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTCCAACCGAAGGCG 1392
             |||||||||||||||||||||||||||||||||||||||||| ||     |||| ||||||
Db        121 TTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCGGCAAGTAAACCAAAGGCG 180

Qy       1393 CAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGAC 1452
              | ||| ||  || ||| |||||| ||| |||||||| |||||   ||||| |||||| |
Db        181 GAAGATAAAAGGGAGGATGAAGAGAGTGCAGGCGTTGGTAACGTCGAAGAAGGCGAAGGC 240

Qy       1453 GAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAA 1512
             ||| |   ||| || ||          || |||||  |  |||||    | |||||| |||
Db        241 GAAGTTTCCGAAGATGA---------AGGCGAAGAAGCCGAAGAAATCGTCGAAGAAGAA 291

Qy       1513 GACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAA 1572
              ||||||||||   ||||| |||   |  |||        |||| |||||| ||||||||
Db        292 CCCGAAGAAGAAGCTGAAGAGGAAGAAGCTGAACCCAAAGAAGTTGAAGAAACCGAAGAA 351

Qy       1573 GAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCC 1632
              ||| |||| |||||||| ||  | |||||||  |  |||||||||  |  |||||||||
Db        352 AAATCGCCGACAGAAGAAAGCGGCAGCGGTTCAAACGCCATCCTGCCTGCCTCGGAAGCC 411

Qy       1633 TCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATT 1692
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATT 471

Qy       1693 CCAAAAAACGGAACGGCGCATTATACCGGCACTTGGGAAGCGCGTATCGGCGTATCGGAT 1752
              |||| || |||||  || || ||||||||||||||||||||||||||||||
```

```
Db        472 CCAAGAACCGGAAAAGCACACTATACCGGCACTTGGGAAGCGCGTATCGGC--------- 522

Qy       1753 AGTGGTACGTCCATTCAAAAGGATAGCTATGCG---AATCAAGGGGCAAAAGCAGAATTT 1809
                 ||  |||||||| ||| |   | |||   ||  ||| ||||||||||||||||
Db        523 ------ACACCCATTCAATGGGACAATCAGGCGGATAAAGAAGCGGCAAAAGCAGAATTT 576

Qy       1810 ACCGTTGATTTCGAAGCGAAGACGGTGTCCGGAATGCTGACAGAAAAAAATGATACAACC 1869
              |||||| ||||||  | |||  || | ||||||| |||||| || ||||| | |  |
Db        577 ACCGTTAATTTCGGCGAGAAATCGATTTCCGGAACGCTGACGGAGAAAAACGGTGTACAA 636

Qy       1870 CCCGCTTTTTATATTGAAAAAGGTGTGATTGACGGTAACGGTTTCCACGCTTTGGCGCAT 1929
              || ||||| |||||||||||  ||   |||||| || |||||||||||||||    || |
Db        637 CCTGCTTTCTATATTGAAAACGGCAAGATTGAGGGCAACGGTTTCCACGCAACAGCACGC 696

Qy       1930 ACTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAA 1989
              ||||| |||||||| ||  | ||||| ||  | |||||||| |||| |  | |||| ||
Db        697 ACTCGTGAGAACGGCATCAATCTTTCGGGAAATGGTTCGACCAACCCCAGAACCTTCCAA 756

Qy       1990 GCCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGT 2049
              ||     |||||| |||| ||| || ||||| |||||||| |||| ||||||||||||
Db        757 GCTAGTGATCTTCGTGTAGAAGGAGGATTTTACGGCCCGCA-GCGGAGGAATTGGGCGGT 815

Qy       2050 AATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGA 2105
              | |||| || | |  ||  ||||||    ||  ||  ||  |  |  || || |
Db        816 ATTATTTTCAATAAGGATGGGAAATCTCTTGGTATAACTGAAGGTACTGAAAATAA 871
```

RESULT 12
AAA81815
ID AAA81815 standard; DNA; 707 BP.
XX
AC AAA81815;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_362 SEQ ID NO:362.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-0103794P.
PR 30-APR-1999; 99US-0132068P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX
PS Claim 7; Page 1606; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent

CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neissariae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 707 BP; 199 A; 155 C; 178 G; 173 T; 0 U; 2 Other;

Query Match 20.2%; Score 428.8; DB 3; Length 707;
Best Local Similarity 77.4%; Pred. No. 2.3e-96;
Matches 546; Conservative 1; Mismatches 149; Indels 9; Gaps 2;

```
Qy         23 GCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATCGGCGGCAATTTCGGCGTGCAGC 82
Db          1 GCATTGTCTTGTTGCCCTTACTTTTAGCATCTTGTATCGGCGGCAATTTCGGCGTGCAGC 60

Qy         83 CTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACTTTCAAGTCTAAGGACGTTCCCA 142
Db         61 CTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACCTTCAAATCTAAGGACGTTCCCA 120

Qy        143 CTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCGGTCAACCGACCCGCCGTTGGTG 202
Db        121 CTCCGCCCCCTGCCGGGTCTTCGGTAGAAACCACGCCGGTCAACCGGCCCGCCGTCGGTG 180

Qy        203 CGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCTGATAAGGATGGCAATGATTTTC 262
Db        181 CGGCAATGCGGCTGCCAAGGCGGAATATTGCTTCCTATAAACAAGACGGTACGGAAATTC 240

Qy        263 CAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAAGAGGAAGATATCCTGTTTTTAT 322
Db        241 CCGACAAGCATCAGGCAGAGGAGCATCTGCCGCTTAAAGAGAAGGATATCCTGTTTTTAG 300

Qy        323 ACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGATAAAATTCGTCAACCAAATCCTA 382
Db        301 ACGGTACGCTGAAAGAACAGGCTGACAAACTTAAAAAGAAAATCAACGAACGGTATTCTG 360

Qy        383 CGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAATATGATTATAAATTTGTAGATG 442
Db        361 ATGTGAGGGTTATCACATCGAAAAAAGAAGAAGAAAAATATCAATATCAATTTGTCCGTG 420

Qy        443 CAGGTTATGTATATACTAAAGACGGAAAAG------ATGAAATTGAGTGGACTTCAAATT 496
Db        421 CGGGCTATGTGTTTACCAGGGCGGAAGGAAAGGATAATGAAAAAGAAAAGACTTCTGATG 480

Qy        497 ACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATATTATTCCGGAGAACATC 556
Db        481 GTAAGGAGTTTGTTAACCGATTTAGTTATGACGGTTTTGTATATTATTCCGGAGAACGTC 540

Qy        557 CTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGCAACTGGCAATATATGA 616
Db        541 CTTCCCAATCTTTACCGAGCGCGGGAACGGTGCAATATTCCGGTAACTGGCAATATATGA 600

Qy        617 CCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGCGAAGATTTGGGTTATA 676
Db        601 CCGATGCCAAACGTCATCGGACANGTAAGGC---GGTTTYCAGTACGGATTTGGGTTATA 657

Qy        677 TCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCGA 721
Db        658 CCACATATTATGGTAATGAAATTGGGGCAACTTCTTATGAGGCTA 702
```

RESULT 13
ABS67377
ID ABS67377 standard; DNA; 3300 BP.
XX
AC ABS67377;
XX
DT 29-NOV-2002 (first entry)
XX
DE Neisseria gonorrhoeae lbpA gene.
XX
KW Gram-negative bacterial bleb; PorB; outer membrane protein;
KW Chlamydia trachomatis infection; Chlamydia pneumoniae infection;
KW protective antigen; antibacterial; vaccine; gene; ds.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200262380-A2.
XX
PD 15-AUG-2002.
XX
PF 08-FEB-2002; 2002WO-EP001356.
XX
PR 08-FEB-2001; 2001GB-00003169.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Berthet FJ, Lobet Y, Poolman J, Verlant VGCL;
XX
DR WPI; 2002-657510/70.
DR P-PSDB; ABG31056.
XX
PT Novel gram-negative bacterial bleb presenting on its surface PorB outer
PT membrane protein from Chlamydia trachomatis or protective antigen from
PT Chlamydia pneumoniae, useful for preventing Chlamydia infection.
XX
PS Disclosure; Page 48-49; 75pp; English.
XX
CC The present invention relates to a new gram-negative bacterial bleb
CC presenting on its surface the PorB outer membrane protein from Chlamydia
CC trachomatis, or a protective antigen from C. pneumoniae. The invention is
CC useful for preventing C. trachomatis or C. pneumoniae infection in a
CC host. The present nucleic acid sequence represents a Neisseria
CC gonorrhoeae gene as described in the invention
XX
SQ Sequence 3300 BP; 897 A; 807 C; 921 G; 675 T; 0 U; 0 Other;

Query Match 4.9%; Score 104.6; DB 6; Length 3300;
Best Local Similarity 74.9%; Pred. No. 2e-15;
Matches 131; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
Qy       1931 CTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAAG 1990
Db          1 CTCGGGATAACGGCATCAATCTTTCGGGAAATGGTTCGACTAATCCTCAAAGTTTCAAAG 60

Qy       1991 CCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGTA 2050
Db         61 CCGACAATCTTCTTGTAACGGGCGGCTTTTACGGCCCGCAGGCGGCGGAATTGGGCGGCA 120

Qy       2051 ATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGA 2105
Db        121 CTATTTTCAATAAGGATGGGAAATCTCTTGGTATAACTGAAGATATTGAAAATGA 175
```

RESULT 14
AAS69547
ID AAS69547 standard; cDNA; 708 BP.
XX
AC AAS69547;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #5351.

XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG05360.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 5351; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 708 BP; 261 A; 127 C; 221 G; 99 T; 0 U; 0 Other;

Query Match 4.6%; Score 98.2; DB 5; Length 708;
Best Local Similarity 65.0%; Pred. No. 4.1e-14;
Matches 145; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy 1396 GATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAA 1455
Db 85 GAAGGAGAAGGGGAAGGGGAAGGGGAAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAA 144

Qy 1456 ATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGAC 1515
Db 145 GAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGAGGAAGAGGAAGAAGAA 204

Qy 1516 GAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAA 1575
Db 205 GAAGAAGAAGAAGAAGAAGAAGAACAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 264

Qy 1576 TTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGC 1618
Db 265 GAAGAAGAAGAAAGAAGAAGCGGCAGCAGCGGCAGCAGCCAGC 307

RESULT 15
AAS75461
ID AAS75461 standard; cDNA; 708 BP.
XX
AC AAS75461;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #11265.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG11274.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 11265; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 708 BP; 261 A; 127 C; 221 G; 99 T; 0 U; 0 Other;

Query Match 4.6%; Score 98.2; DB 5; Length 708;
Best Local Similarity 65.0%; Pred. No. 4.1e-14;
Matches 145; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy 1396 GATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAA 1455
Db 85 GAAGGAGAAGGGGAAGGGGAAGGGGAAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAA 144

Qy 1456 ATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGAC 1515

```
Db        145 GAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGAGGAAGAGGAAGAAGAA 204

Qy       1516 GAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAA 1575

Db        205 GAAGAAGAAGAAGAAGAAGAAGAACAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 264

Qy       1576 TTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGC 1618

Db        265 GAAGAAGAAGAAAGAAGAAGCGGCAGCAGCGGCAGCAGCCAGC 307
```

Search completed: August 24, 2004, 14:38:37
Job time : 728.611 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 12:38:10 ; Search time 154.323 Seconds
(without alignments)
7637.968 Million cell updates/sec

Title: US-10-735-098-9
Perfect score: 2124
Sequence: 1 atgtgtaaaccgaattatgg.....acaaggaggcaacacgatga 2124

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	86.2	4.1	7218	1	US-08-232-463-14	Sequence 14, Appl
c 2	86	4.0	929	4	US-09-671-317-14	Sequence 14, Appl
c 3	86	4.0	1001	4	US-09-671-317-439	Sequence 439, App
4	79.8	3.8	43795	3	US-08-742-185-101	Sequence 101, App
5	75.8	3.6	2223	1	US-08-257-073-4	Sequence 4, Appli
6	74.8	3.5	3211	2	US-08-574-959A-8	Sequence 8, Appli
7	74.8	3.5	3211	3	US-09-357-014-8	Sequence 8, Appli
8	74.8	3.5	3901	2	US-08-574-959A-6	Sequence 6, Appli
9	74.8	3.5	3901	3	US-09-357-014-6	Sequence 6, Appli
c 10	74.4	3.5	16442	3	US-08-781-891-208	Sequence 208, App
c 11	74.4	3.5	16442	4	US-09-618-166-208	Sequence 208, App
12	71.4	3.4	390	3	US-09-197-649-7	Sequence 7, Appli
13	69.4	3.3	1276	3	US-09-177-325-2	Sequence 2, Appli
14	69.4	3.3	1276	3	US-09-411-812A-2	Sequence 2, Appli
15	69.4	3.3	1276	4	US-09-590-113-2	Sequence 2, Appli
16	69.2	3.3	3337	1	US-08-072-610-1	Sequence 1, Appli
17	69.2	3.3	3337	2	US-08-719-822B-1	Sequence 1, Appli
18	69.2	3.3	3337	3	US-09-092-458-1	Sequence 1, Appli
19	68.4	3.2	966	2	US-08-766-738-2	Sequence 2, Appli
20	68.4	3.2	966	4	US-09-262-610-2	Sequence 2, Appli
21	67.2	3.2	1236	2	US-08-741-134-5	Sequence 5, Appli
22	65.4	3.1	5394	3	US-08-688-376-1	Sequence 1, Appli
23	65	3.1	3489	2	US-08-728-323A-1	Sequence 1, Appli
24	65	3.1	3489	4	US-09-298-568-1	Sequence 1, Appli
25	65	3.1	3489	4	US-09-410-399-1	Sequence 1, Appli
c 26	65	3.1	32207	2	US-08-770-379-20	Sequence 20, Appl
c 27	65	3.1	32207	3	US-08-757-669A-20	Sequence 20, Appl
c 28	65	3.1	32207	4	US-09-230-371A-20	Sequence 20, Appl
29	64.8	3.1	2230	3	US-08-448-194-7	Sequence 7, Appli
30	64.8	3.1	2230	4	US-08-867-921-7	Sequence 7, Appli
31	64.4	3.0	2106	3	US-08-613-009A-4	Sequence 4, Appli
32	64.4	3.0	2106	4	US-08-778-570B-4	Sequence 4, Appli
33	64.4	3.0	2106	4	US-09-059-584-4	Sequence 4, Appli
34	64.4	3.0	2247	3	US-08-613-009A-3	Sequence 3, Appli
35	64.4	3.0	2247	4	US-08-778-570B-3	Sequence 3, Appli
36	64.4	3.0	2247	4	US-09-059-584-3	Sequence 3, Appli
37	64.4	3.0	8266	4	US-09-059-584-54	Sequence 54, Appl
38	62.8	3.0	3955	4	US-09-976-594-207	Sequence 207, App
39	61.6	2.9	1052	1	US-08-466-603-1	Sequence 1, Appli
40	61.6	2.9	1052	1	US-08-314-503A-1	Sequence 1, Appli
41	61.6	2.9	1052	1	US-08-468-066-1	Sequence 1, Appli
42	61.6	2.9	1052	2	US-08-466-717-1	Sequence 1, Appli
43	61.6	2.9	1052	3	US-08-466-743-1	Sequence 1, Appli
44	61.6	2.9	1052	5	PCT-US95-12414-1	Sequence 1, Appli
45	61.2	2.9	2139	4	US-09-059-584-50	Sequence 50, Appl

ALIGNMENTS

```
RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

 Query Match 4.1%;  Score 86.2;  DB 1;  Length 7218;
```

```
 Best Local Similarity   4.0%;  Pred. No. 2.7e-12;
 Matches   16;  Conservative  249;  Mismatches  132;  Indels    0;  Gaps    0;

Qy       1331 TTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTCCAACCGAAGG 1390
             ||| |  | : : ::::   ::: ::: :::::  ::: :    :  :   ::  ::::: 
Db       1442 TTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1383

Qy       1391 CGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAG 1450
              : :::: :::::::::: :::::::: : ::: :  :: :: :: :::::::: :::: 
Db       1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323

Qy       1451 ACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATG 1510
             : ::::  :: :: :::::::: :  ::::: ::::  : :::::: ::::: ::::: : 
Db       1322 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263

Qy       1511 AAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAG 1570
             :::: ::::::::::  :::::::::  ::::::::::  :::::::::::   :::: 
Db       1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203

Qy       1571 AAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAG 1630
             :::::  :  :: ::::::::: :: :: ::  ::: :: :    :   :    ::::: 
Db       1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143

Qy       1631 CCTCTAAAGGCAGGGACATCGACCTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACA 1690
                 ::::: ::::: :  ::         :::::: :   : : :: ::::  :: : 
Db       1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1083

Qy       1691 TTCCAAAAAACGGAACGGCGCATTATACCGGCACTTG 1727
                ::::: ::::   ||||   | ||   || ||
Db       1082 RRRRRRRRRRRRRRRRATCGCAAGCTCCCTCGACCTG 1046


RESULT 2
US-09-671-317-14/c
; Sequence 14, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
;  FILE REFERENCE:  62.US3.CIP
;  CURRENT APPLICATION NUMBER: US/09/671,317
;  CURRENT FILING DATE:  2000-09-27
;  PRIOR APPLICATION NUMBER:  US 09/536,178
;  PRIOR FILING DATE:  2000-03-23
;  PRIOR APPLICATION NUMBER:  PCT/IB00/00403
;  PRIOR FILING DATE:  2000-03-24
;  PRIOR APPLICATION NUMBER:  US 60/126,269
;  PRIOR FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 60/131,961
;  PRIOR FILING DATE:  1999-04-30
;  NUMBER OF SEQ ID NOS:  977
;  SOFTWARE:  Patent.pm
; SEQ ID NO 14
;   LENGTH: 929
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 501
;   OTHER INFORMATION: 12-454-363 : polymorphic base A or G
;   NAME/KEY: misc_binding
;   LOCATION: 481..500
;   OTHER INFORMATION: 12-454-363.mis1, potential
;   NAME/KEY: misc_binding
;   LOCATION: 502..521
;   OTHER INFORMATION: 12-454-363.mis2, potential complement
;   NAME/KEY: primer_bind
;   LOCATION: 139..158
;   OTHER INFORMATION: upstream amplification primer
;   NAME/KEY: primer_bind
;   LOCATION: 634..652
;   OTHER INFORMATION: downstream amplification primer, complement
;   NAME/KEY: misc_binding
;   LOCATION: 489..513
;   OTHER INFORMATION: 12-454-363 potential probe
;   NAME/KEY: misc_feature
;   LOCATION: 674..679,881..882,892..893
;   OTHER INFORMATION: n=a, g, c or t
US-09-671-317-14

  Query Match             4.0%;  Score 86;  DB 4;  Length 929;
  Best Local Similarity   64.4%;  Pred. No. 1.1e-12;
  Matches  139;  Conservative    0;  Mismatches   76;  Indels    1;  Gaps    1;

Qy       1380 CCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAA 1439
             ||| |  || |   | || ||||| |  || ||||| |  | ||  |  ||  | |   |
Db        880 CCATCTCAAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 821

Qy       1440 AGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGA 1499
             ||||   ||||| |||   |  || ||||||      | ||| ||||  | |||||| ||
Db        820 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 761

Qy       1500 AGACGAAGATGAAGACGAAGAA-GAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGG 1558
             ||| ||||| ||||| |||||| |||   |||||||||   |||||||||  ||||| | 
Db        760 AGAAGAAGAAGAAGAAGAAGAAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 701

Qy       1559 AAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCA 1594
             |||||   |||||||||          ||||| | |
Db        700 AAGAAGAAGAAGAAGAAGAAGNNNNNNAAGAAAGAA 665


RESULT 3
US-09-671-317-439/c
; Sequence 439, Application US/09671317
; Patent No. 6528260
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
;  FILE REFERENCE:  62.US3.CIP
;  CURRENT APPLICATION NUMBER: US/09/671,317
;  CURRENT FILING DATE:  2000-09-27
;  PRIOR APPLICATION NUMBER:  US 09/536,178
;  PRIOR FILING DATE:  2000-03-23
;  PRIOR APPLICATION NUMBER:  PCT/IB00/00403
;  PRIOR FILING DATE:  2000-03-24
;  PRIOR APPLICATION NUMBER:  US 60/126,269
;  PRIOR FILING DATE:  1999-03-25
;  PRIOR APPLICATION NUMBER:  US 60/131,961
;  PRIOR FILING DATE:  1999-04-30
;  NUMBER OF SEQ ID NOS:  977
;  SOFTWARE:  Patent.pm
; SEQ ID NO 439
;   LENGTH: 1001
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 501
;   OTHER INFORMATION: 12-454-242 : deletion  AT
;   NAME/KEY: misc_binding
;   LOCATION: 481..500
;   OTHER INFORMATION: 12-454-242.mis1, potential
;   NAME/KEY: primer_bind
;   LOCATION: 260..279
;   OTHER INFORMATION: upstream amplification primer
;   NAME/KEY: primer_bind
;   LOCATION: 755..773
;   OTHER INFORMATION: downstream amplification primer, complement
;   NAME/KEY: misc_feature
;   LOCATION: 795..800
```

```
;   OTHER INFORMATION: n=a, g, c or t
US-09-671-317-439

  Query Match             4.0%;  Score 86;  DB 4;  Length 1001;
  Best Local Similarity   64.4%;  Pred. No. 1.2e-12;
  Matches  139;  Conservative    0;  Mismatches   76;  Indels    1;  Gaps    1;

Qy       1380 CCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAA 1439
             ||| |  || |   | || |||||  |  || |||||  |  || |  ||  | |   |
Db       1001 CCATCTCAAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 942

Qy       1440 AGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGA 1499
              ||||  |||||  |||   |  || ||||||      | |||  ||||  | ||||||  ||
Db        941 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 882

Qy       1500 AGACGAAGATGAAGACGAAGAA-GAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGG 1558
              ||| |||||  |||||  ||||||  |||   |||||||||   ||||||||||  |||||  |
Db        881 AGAAGAAGAAGAAGAAGAAGAAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 822

Qy       1559 AAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCA 1594
              |||||   |||||||||           |||||  | |
Db        821 AAGAAGAAGAAGAAGAAGAAGNNNNNNAAGAAAGAA 786


RESULT 4
US-08-742-185-101
; Sequence 101, Application US/08742185
; Patent No. 6020476
;  GENERAL INFORMATION:
;    APPLICANT:  Page, David C.
;    APPLICANT:  Reijo, Renee
;    APPLICANT:  Saxena, Richa
;    APPLICANT:  Hawkins,Trevor
;    APPLICANT:  Reeve, Mary Pat
;    TITLE OF INVENTION:  DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA
;    NUMBER OF SEQUENCES:  102
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  US
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,185
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/690,734
;      FILING DATE:  31-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/310,429
;      FILING DATE:  22-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  WHI94-07A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  101:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43795 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-742-185-101

  Query Match             3.8%;  Score 79.8;  DB 3;  Length 43795;
  Best Local Similarity   64.2%;  Pred. No. 3e-10;
  Matches  120;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy       1388 AGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCG 1447
             | | | |||| || ||||| |  | ||  || | ||| |  ||    |  || | |
Db      38855 AAGAGGAGGAGGAGGAGGGAGGGGGAGGAGGAGGAGGAGGAGAATGAGAAGAAAAGAAAA 38914

Qy       1448 AAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAG 1507
              |||| |||   |  || ||||||  |     | ||| ||   | ||  || || || ||||
Db      38915 AAGAAGAAGAAGAAGAGGAAGAAGGGGAAGAAGAAGAGAAGGAAGCGGAAGAGGAAGAAG 38974

Qy       1508 ATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCG 1567
              | ||||  |||||  |||   |||||||||  | ||||||||||  ||  ||||||||||   |
Db      38975 AAGAAGGAGAAGAGGAAGAAGAAGAAGAAGCAGAAGAAGAAGAAGAGGAGGAAGAAGAAG 39034

Qy       1568 AAGAAGA 1574
              |||||||
Db      39035 AAGAAGA 39041


RESULT 5
US-08-257-073-4
; Sequence 4, Application US/08257073
; Patent No. 5766597
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  de Taisne, Charles
;    APPLICANT:  Tine, John A.
;    TITLE OF INVENTION:  MALARIA RECOMBINANT POXVIRUS VACCINE
;    NUMBER OF SEQUENCES:  143
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford, P.C.
;      STREET:  530 Fifth Avenue, 25th Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/257,073
;      FILING DATE:  09-JUN-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/075,783
;      FILING DATE:  11-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/852,305
;      FILING DATE:  18-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/672,183
;      FILING DATE:  20-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2570
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;      TELEX:  425066 CURTMS
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2223 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-257-073-4
```

```
  Query Match             3.6%;  Score 75.8;  DB 1;  Length 2223;
  Best Local Similarity   58.0%;  Pred. No. 7.7e-10;
  Matches  134;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy       1357 ATAAAAACCGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAG 1416
             ||| || | |||        ||| ||  ||| ||      |||| ||||| |   | ||| |
Db       1984 ATAGAAGCTGAAGTGGATGCCCTCGCACCAAAAAATAAGGAAGAAGAAGAAAAAGAAAAA 2043

Qy       1417 GGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACC 1476
             |    ||     || || |   || ||    |||| ||||  |   | ||| ||
Db       2044 GAAAAAGAAAAGGAAAAAGAAGAAAAAGAAAAAGAAGAAAAAGAAAAAGAAAAAGAAGAA 2103

Qy       1477 GGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAA 1536
                |||  |||    |||| | |||||  |||| |||||  || ||||||  |||||||||
Db       2104 AAAGAAAAAGAAAAAGAAAAAGAAGAAAAAGAAGAAGAAAAAAAAGAAAAAGAAGAAGAA 2163

Qy       1537 CCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAA 1587
             | ||||||||||  |||||   || ||| ||| |          | | ||||
Db       2164 CAAGAAGAAGAAGAAGAAGAAATAGTACCAGAAAATTTGACAACTGAAGAA 2214


RESULT 6
US-08-574-959A-8
; Sequence 8, Application US/08574959A
; Patent No. 5962224
;  GENERAL INFORMATION:
;    APPLICANT:  Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;    APPLICANT:  and Jack L. Strominger
;    TITLE OF INVENTION:  p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/574,959A
;      FILING DATE:  19-DEC-95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3211 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  439..3157
US-08-574-959A-8

  Query Match             3.5%;  Score 74.8;  DB 2;  Length 3211;
  Best Local Similarity   61.1%;  Pred. No. 1.7e-09;
  Matches  121;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
             |||| ||||| |  || ||||| | | |  | ||  | |  |||||  |||||| |
Db       2423 AGGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAGGAAGAAGAGGAAGAGG 2482

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
             |    | |  ||| |||        |  || ||||  | | |   ||||| ||||| || |
Db       2483 AGGAAGACTTTGAGGAAGAGGAAGAGGATGAAGAGGAATATTTTGAAGAGGAAGAAGAGG 2542

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
             | ||||||||  | ||  ||||||  ||| |||||||  |  |||  |||   |||||||
Db       2543 AGGAAGAAGAGTTTGAGGAAGAATTTGAGGAAGAAGAAGGTGAGTTAGAGGAAGAAGAAG 2602

Qy       1574 AATTGCCGGCAGAAGAAG 1591
             ||  |  ||  || || | |
Db       2603 AAGAGGAGGATGAGGAGG 2620


RESULT 7
US-09-357-014-8
; Sequence 8, Application US/09357014
; Patent No. 6291645
;  GENERAL INFORMATION:
;       APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;                  and Jack L. Strominger
;       TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;                           AND USES THEREFOR
;       NUMBER OF SEQUENCES: 22
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 60 State Street, Suite 510
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/357,014
;            FILING DATE: 19-Jul-1999
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/574,959
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mandragouras, Amy E.
;            REGISTRATION NUMBER: 36,207
;            REFERENCE/DOCKET NUMBER: DFN-008
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)227-5941
;  INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 3211 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       FEATURE:
;            NAME/KEY: CDS
;            LOCATION: 439..3157
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-357-014-8

  Query Match             3.5%;  Score 74.8;  DB 3;  Length 3211;
  Best Local Similarity   61.1%;  Pred. No. 1.7e-09;
  Matches  121;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
             |||| ||||| |  || ||||| | | |  | ||  | |  |||||  |||||| |
Db       2423 AGGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAGGAAGAAGAGGAAGAGG 2482

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
             |    | |  ||| |||        |  || ||||  | | |   ||||| ||||| || |
Db       2483 AGGAAGACTTTGAGGAAGAGGAAGAGGATGAAGAGGAATATTTTGAAGAGGAAGAAGAGG 2542
```

```
Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
             |||||||||  | || ||||||  ||| |||||||  |  |||  |||     |||||||
Db       2543 AGGAAGAAGAGTTTGAGGAAGAATTTGAGGAAGAAGAAGGTGAGTTAGAGGAAGAAGAAG 2602

Qy       1574 AATTGCCGGCAGAAGAAG 1591
              ||  |  ||  || || |
Db       2603 AAGAGGAGGATGAGGAGG 2620


RESULT 8
US-08-574-959A-6
; Sequence 6, Application US/08574959A
; Patent No. 5962224
;  GENERAL INFORMATION:
;    APPLICANT:  Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;    APPLICANT:  and Jack L. Strominger
;    TITLE OF INVENTION:  p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/574,959A
;      FILING DATE:  19-DEC-95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3901 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  439..3847
US-08-574-959A-6

  Query Match             3.5%;  Score 74.8;  DB 2;  Length 3901;
  Best Local Similarity   61.1%;  Pred. No. 1.8e-09;
  Matches  121;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
             |||| |||||  |  || |||||  |  |  |  |  ||  | |  |||||  |||||  |
Db       3113 AGGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAGGAAGAAGAGGAAGAGG 3172

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
              |    | |  ||| |||        |  || ||||  | | |  ||||||  |||||  || |
Db       3173 AGGAAGACTTTGAGGAAGAGGAAGAGGATGAAGAGGAATATTTTGAAGAGGAAGAAGAGG 3232

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
              | |||||||||  | || ||||||  ||| |||||||  |  |||  |||     |||||||
Db       3233 AGGAAGAAGAGTTTGAGGAAGAATTTGAGGAAGAAGAAGGTGAGTTAGAGGAAGAAGAAG 3292

Qy       1574 AATTGCCGGCAGAAGAAG 1591
              ||  |  ||  || || |
Db       3293 AAGAGGAGGATGAGGAGG 3310


RESULT 9
US-09-357-014-6
; Sequence 6, Application US/09357014
; Patent No. 6291645
;  GENERAL INFORMATION:
;       APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;                  and Jack L. Strominger
;       TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;                           AND USES THEREFOR
;       NUMBER OF SEQUENCES: 22
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 60 State Street, Suite 510
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/357,014
;            FILING DATE: 19-Jul-1999
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/574,959
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mandragouras, Amy E.
;            REGISTRATION NUMBER: 36,207
;            REFERENCE/DOCKET NUMBER: DFN-008
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)227-5941
;  INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 3901 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  439..3847
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-357-014-6

  Query Match             3.5%;  Score 74.8;  DB 3;  Length 3901;
  Best Local Similarity   61.1%;  Pred. No. 1.8e-09;
  Matches  121;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
             |||| |||||  |  || |||||  |  |  |  |  ||  | |  |||||  |||||  |
Db       3113 AGGAAGAAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAGGAAGAAGAGGAAGAGG 3172

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
              |    | |  ||| |||        |  || ||||  | | |  ||||||  |||||  || |
Db       3173 AGGAAGACTTTGAGGAAGAGGAAGAGGATGAAGAGGAATATTTTGAAGAGGAAGAAGAGG 3232

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
              | |||||||||  | || ||||||  ||| |||||||  |  |||  |||     |||||||
Db       3233 AGGAAGAAGAGTTTGAGGAAGAATTTGAGGAAGAAGAAGGTGAGTTAGAGGAAGAAGAAG 3292

Qy       1574 AATTGCCGGCAGAAGAAG 1591
              ||  |  ||  || || |
Db       3293 AAGAGGAGGATGAGGAGG 3310


RESULT 10
US-08-781-891-208/c
; Sequence 208, Application US/08781891
```

```
; Patent No. 6090620
;  GENERAL INFORMATION:
;    APPLICANT:  Fu, Ying-Hui
;    APPLICANT:  Yu, Chang-En
;    APPLICANT:  Oshima, Junko
;    APPLICANT:  Mulligan, John T.
;    APPLICANT:  Schellenberg, Gerald D.
;    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
;    TITLE OF INVENTION:  WERNER'S SYNDROME
;    NUMBER OF SEQUENCES:  209
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,891
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6090620tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,317
;      REFERENCE/DOCKET NUMBER:  240052.419
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  208:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16442 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-781-891-208

  Query Match             3.5%;  Score 74.4;  DB 3;  Length 16442;
  Best Local Similarity   65.1%;  Pred. No. 4.7e-09;
  Matches  125;  Conservative    0;  Mismatches   66;  Indels    1;  Gaps    1;

Qy       1386 GAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAG 1445
             | ||| |||||| || |||  ||| | ||  || | ||| | |    |    || | |
Db      16298 GGAGGAGCAGGAGGAGGAGCAGGA-GCAGGAGGAGGAGGAGCAGGAGGAGCAGGAGGAGG 16240

Qy       1446 CGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGA 1505
             | ||  | |   || |  | || | |    ||||  | ||  | ||||| ||||| ||
Db      16239 AGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGTAGAAGGAGAAGAAGAAGAAGA 16180

Qy       1506 AGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACC 1565
             ||| ||||| |||||||||   |||||||||  ||||||||||  ||||| ||||||
Db      16179 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 16120

Qy       1566 CGAAGAAGAATT 1577
              |||||||||  |
Db      16119 AGAAGAAGAAAT 16108


RESULT 11
US-09-618-166-208/c
; Sequence 208, Application US/09618166
; Patent No. 6583112
;   GENERAL INFORMATION:
;        APPLICANT: Fu, Ying-Hui
;                   Yu, Chang-En
;                   Oshima, Junko
;                   Mulligan, John T.
;                   Schellenberg, Gerald D.
;        TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
;                            WERNER'S SYNDROME
;        NUMBER OF SEQUENCES: 209
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/618,166
;             FILING DATE: 17-Jul-2000
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 240052.419C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 208:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 16442 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 208:
US-09-618-166-208

  Query Match             3.5%;  Score 74.4;  DB 4;  Length 16442;
  Best Local Similarity   65.1%;  Pred. No. 4.7e-09;
  Matches  125;  Conservative    0;  Mismatches   66;  Indels    1;  Gaps    1;

Qy       1386 GAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAG 1445
             | ||| |||||| || |||  ||| | ||  || | ||| | |    |    || | |
Db      16298 GGAGGAGCAGGAGGAGGAGCAGGA-GCAGGAGGAGGAGGAGCAGGAGGAGCAGGAGGAGG 16240

Qy       1446 CGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGA 1505
             | ||  | |   || |  | || | |    ||||  | ||  | ||||| ||||| ||
Db      16239 AGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGTAGAAGGAGAAGAAGAAGAAGA 16180

Qy       1506 AGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACC 1565
             ||| ||||| |||||||||   |||||||||  ||||||||||  ||||| ||||||
Db      16179 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 16120

Qy       1566 CGAAGAAGAATT 1577
              |||||||||  |
Db      16119 AGAAGAAGAAAT 16108


RESULT 12
US-09-197-649-7
; Sequence 7, Application US/09197649
; Patent No. 6194550
; GENERAL INFORMATION:
;  APPLICANT: Gold, Larry
;  APPLICANT: Tuerk, Craig
;  APPLICANT: Pribnow, David
;  APPLICANT: Smith, Jonathan D.
;  TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
;  FILE REFERENCE: NEX02/C1-CON
;  CURRENT APPLICATION NUMBER: US/09/197,649
;  CURRENT FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: 07/829,461
;  EARLIER FILING DATE: 1992-01-31
;  EARLIER APPLICATION NUMBER: 07/739,055
;  EARLIER FILING DATE: 1991-08-01
;  EARLIER APPLICATION NUMBER: 07/561,968
```

; EARLIER FILING DATE: 1990-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
; OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-197-649-7

Query Match 3.4%; Score 71.4; DB 3; Length 390;
Best Local Similarity 54.9%; Pred. No. 4.6e-09;
Matches 141; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

```
Qy       1365 CGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGG 1424
             ||| | || || | |   | | | || || || | ||||| || |  | |
Db        125 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA 184

Qy       1425 CGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGA 1484
             ||  ||  |||   | ||  ||| |||||   || ||| || ||   |  ||  |||||
Db        185 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA 244

Qy       1485 AGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGA 1544
             | ||  || || || ||||| || ||  ||||| || ||    ||| || ||    || ||
Db        245 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA 304

Qy       1545 AGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAACGGCGGTTC 1604
             || |  || || || ||  ||| || ||      |  || || | | ||| ||
Db        305 CGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA 364

Qy       1605 AGGCAGCATCCTGCCCA 1621
                | |  |||  |  | |
Db        365 CGACGCCATGGTTGCGA 381
```

RESULT 13
US-09-177-325-2
; Sequence 2, Application US/09177325B
; Patent No. 6214983
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sakin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, A
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: LUD 5525
; CURRENT APPLICATION NUMBER: US/09/177,325B
; CURRENT FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 2
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-177-325-2

Query Match 3.3%; Score 69.4; DB 3; Length 1276;
Best Local Similarity 56.3%; Pred. No. 2.7e-08;
Matches 130; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```
Qy       1359 AAAAACCGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGG 1418
             | ||| ||| |   | || |  ||   |  |    || || || | ||| ||||||
Db        701 AGAAGCCGCAGCAGCAGCAGCAGAAGTAATAGTAGTAGAAGACGAAGAGGAGGAAGAGAA 760

Qy       1419 TGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGG 1478
             |  |  |  ||  |    ||||    ||||| |||   |  || |||| | |   |
Db        761 GGAGGAGGAGGAGGAGAAAGAAGAGGAGGAAGAAGAAGGAGAAGAAGAAGGAGGAGGAGA 820

Qy       1479 AGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACC 1538
             ||| | ||  | || ||  | ||  | ||| | ||| |||  ||||   || |||||
Db        821 AGAAGGAGAAGAAGGAGGAGGAGGAGGAGAAGGAGAAGAAACAGAAGAAGAGGAAGAGGA 880

Qy       1539 TGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGA 1589
             |||||||||| ||||||||||||||       ||| ||| |  | | || |
Db        881 AGAAGAAGAAGAGGAAGAGGAAGAACAAATTAAAGCATTTCAAGAAAAACA 931
```

RESULT 14
US-09-411-812A-2
; Sequence 2, Application US/09411812A
; Patent No. 6261778
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, And
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: LUD 5525.1 CIP
; CURRENT APPLICATION NUMBER: US/09/411,812A
; CURRENT FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 09/177,325
; PRIOR FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-411-812A-2

Query Match 3.3%; Score 69.4; DB 3; Length 1276;
Best Local Similarity 56.3%; Pred. No. 2.7e-08;
Matches 130; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```
Qy       1359 AAAAACCGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGG 1418
             | ||| ||| |   | || |  ||   |  |    || || || | ||| ||||||
Db        701 AGAAGCCGCAGCAGCAGCAGCAGAAGTAATAGTAGTAGAAGACGAAGAGGAGGAAGAGAA 760

Qy       1419 TGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGG 1478
             |  |  |  ||  |    ||||    ||||| |||   |  || |||| | |   |
Db        761 GGAGGAGGAGGAGGAGAAAGAAGAGGAGGAAGAAGAAGGAGAAGAAGAAGGAGGAGGAGA 820

Qy       1479 AGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACC 1538
             ||| | ||  | || ||  | ||  | ||| | ||| |||  ||||   || |||||
Db        821 AGAAGGAGAAGAAGGAGGAGGAGGAGGAGAAGGAGAAGAAACAGAAGAAGAGGAAGAGGA 880

Qy       1539 TGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGA 1589
             |||||||||| ||||||||||||||       ||| ||| |  | | || |
Db        881 AGAAGAAGAAGAGGAAGAGGAAGAACAAATTAAAGCATTTCAAGAAAAACA 931
```

RESULT 15
US-09-590-113-2
; Sequence 2, Application US/09590113
; Patent No. 6306389
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sakin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, And U
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: LUD 5525
; CURRENT APPLICATION NUMBER: US/09/590,113
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: US 09/177,325
; PRIOR FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 2
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-590-113-2

Query Match 3.3%; Score 69.4; DB 4; Length 1276;
Best Local Similarity 56.3%; Pred. No. 2.7e-08;
Matches 130; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```
Qy       1359 AAAAACCGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGG 1418
              |  ||  |||  |     |  ||  |    ||    |   |      ||  ||  ||  |  |||  ||||||
Db        701 AGAAGCCGCAGCAGCAGCAGCAGAAGTAATAGTAGTAGAAGACGAAGAGGAGGAAGAGAA 760

Qy       1419 TGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGG 1478
                |   |   |   ||   |       ||||     |||||  |||    |   ||  ||||  |  |     |
Db        761 GGAGGAGGAGGAGGAGAAAGAAGAGGAGGAAGAAGAAGGAGAAGAAGAAGGAGGAGGAGA 820

Qy       1479 AGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACC 1538
              |||  |  ||   |  ||  ||   |  ||   |  |||  |  |||  |||   ||||   ||  |||||
Db        821 AGAAGGAGAAGAAGGAGGAGGAGGAGGAGAAGGAGAAGAAACAGAAGAAGAGGAAGAGGA 880

Qy       1539 TGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGA 1589
                |||||||||||  ||||||||||||||       |||  |||  |   |  |  ||  |
Db        881 AGAAGAAGAAGAGGAAGAGGAAGAACAAATTAAAGCATTTCAAGAAAAACA 931
```

Search completed: August 25, 2004, 05:32:08
Job time : 155.323 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 23:48:06 ; Search time 1069.14 Seconds
(without alignments)
9777.159 Million cell updates/sec

Title: US-10-735-098-9
Perfect score: 2124
Sequence: 1 atgtgtaaaccgaattatgg.....acaaggaggcaacacgatga 2124

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3237270 seqs, 2460713050 residues

Total number of hits satisfying chosen parameters: 6474540

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

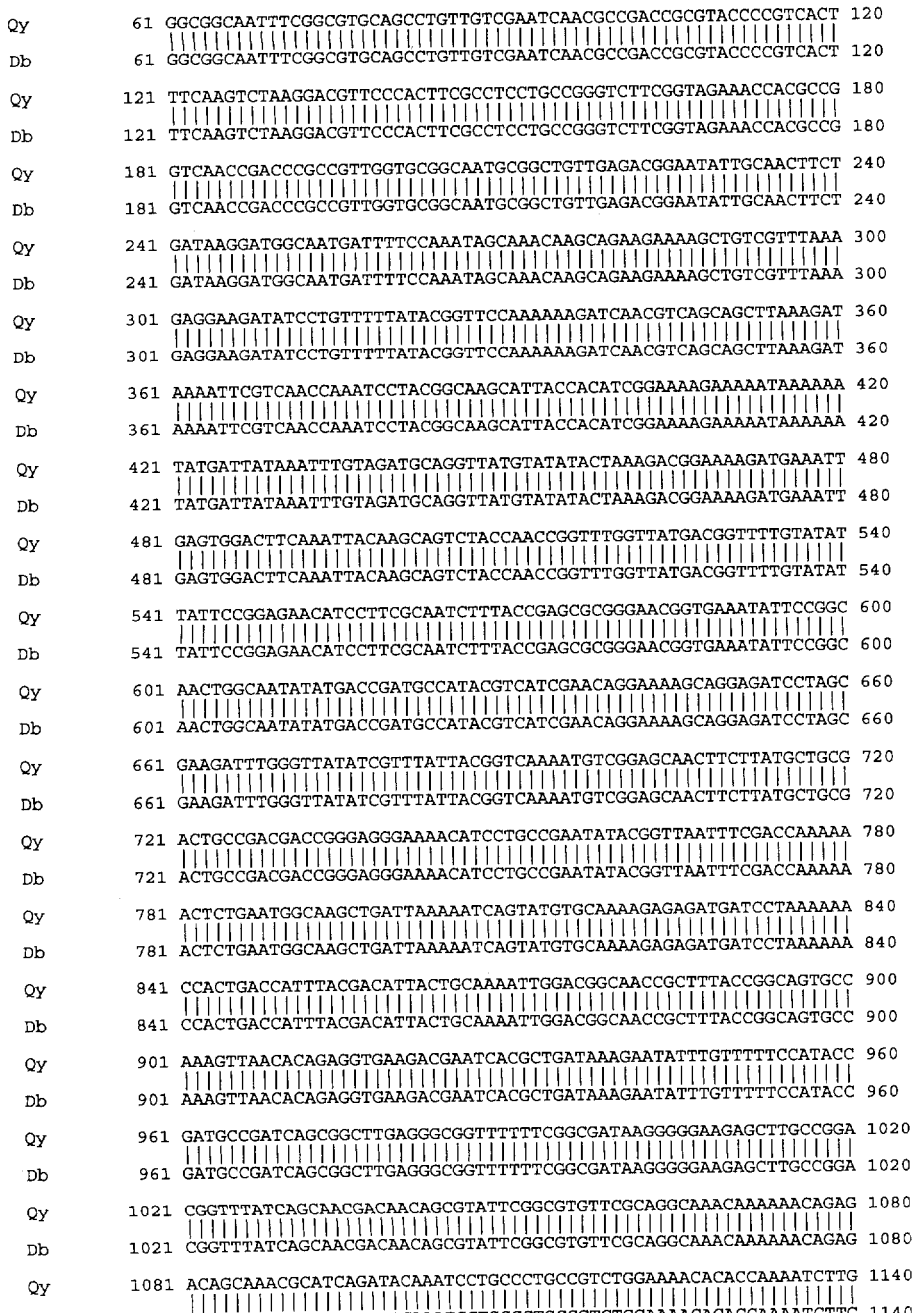
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2124	100.0	2124	17	US-10-735-098-9	Sequence 9, Appli
	2	1562.4	73.6	2169	17	US-10-735-098-3	Sequence 3, Appli
	3	1516.8	71.4	2226	13	US-10-282-122A-29815	Sequence 29815, A
	4	1499.8	70.6	2226	17	US-10-735-098-5	Sequence 5, Appli
	5	1296.2	61.0	2277	17	US-10-735-098-1	Sequence 1, Appli
	6	1291.8	60.8	2262	17	US-10-735-098-7	Sequence 7, Appli
	7	497.6	23.4	1000	17	US-10-343-561-15	Sequence 15, Appl
	8	104.6	4.9	3300	17	US-10-467-534-80	Sequence 80, Appl
c	9	98.4	4.6	374849	13	US-10-087-192-1627	Sequence 1627, Ap
c	10	97.2	4.6	31124	13	US-10-087-192-463	Sequence 463, App
	11	96.6	4.5	37265	13	US-10-087-192-49	Sequence 49, Appl
c	12	96	4.5	115223	17	US-10-322-281-773	Sequence 773, App
c	13	95.4	4.5	635	13	US-10-027-632-269927	Sequence 269927,
c	14	95.4	4.5	635	16	US-10-027-632-269927	Sequence 269927,
c	15	93.8	4.4	305	9	US-09-864-761-19262	Sequence 19262, A
c	16	93.8	4.4	496	9	US-09-864-761-2534	Sequence 2534, Ap
c	17	93.6	4.4	96602	16	US-10-085-117-61	Sequence 61, Appl
c	18	90.6	4.3	39443	16	US-10-085-117-313	Sequence 313, App
c	19	90.2	4.2	276	9	US-09-864-761-20595	Sequence 20595, A
c	20	89.6	4.2	32069	15	US-10-004-113-7	Sequence 7, Appli
	21	89	4.2	48652	13	US-10-087-192-859	Sequence 859, App
	22	88.6	4.2	96596	12	US-10-052-482-70	Sequence 70, Appl
c	23	88.6	4.2	96597	12	US-10-052-482-103	Sequence 103, App
	24	88.4	4.2	522	14	US-10-101-487-71	Sequence 71, Appl
c	25	88.4	4.2	530	14	US-10-101-487-73	Sequence 73, Appl
	26	88.4	4.2	554	14	US-10-101-487-69	Sequence 69, Appl
	27	88.4	4.2	554	14	US-10-101-487-106	Sequence 106, App
c	28	88	4.1	54786	12	US-10-052-482-211	Sequence 211, App
	29	87.2	4.1	225883	15	US-10-175-523-57	Sequence 57, Appl
c	30	86.2	4.1	193853	13	US-10-087-192-1663	Sequence 1663, Ap
c	31	86	4.0	929	13	US-10-294-934-14	Sequence 14, Appl
c	32	86	4.0	1001	13	US-10-294-934-439	Sequence 439, App
c	33	86	4.0	143899	10	US-09-972-546-15	Sequence 15, Appl
	34	85.6	4.0	4316	9	US-09-880-107-3713	Sequence 3713, Ap
	35	85.6	4.0	4316	13	US-10-257-021-81	Sequence 81, Appl
c	36	85	4.0	559	9	US-09-864-761-7684	Sequence 7684, Ap
c	37	85	4.0	210528	13	US-10-087-192-289	Sequence 289, App
	38	84.2	4.0	648	12	US-10-152-319A-2106	Sequence 2106, Ap
c	39	84	4.0	286	15	US-10-029-386-18943	Sequence 18943, A
c	40	84	4.0	577	15	US-10-029-386-5187	Sequence 5187, Ap
c	41	84	4.0	138115	17	US-10-322-281-377	Sequence 377, App
	42	83.6	3.9	536	15	US-10-029-386-19974	Sequence 19974, A
	43	83.6	3.9	599	15	US-10-029-386-6243	Sequence 6243, Ap
c	44	83.6	3.9	33454	13	US-10-087-192-1111	Sequence 1111, Ap
	45	83.2	3.9	6944	15	US-10-172-086-111	Sequence 111, App

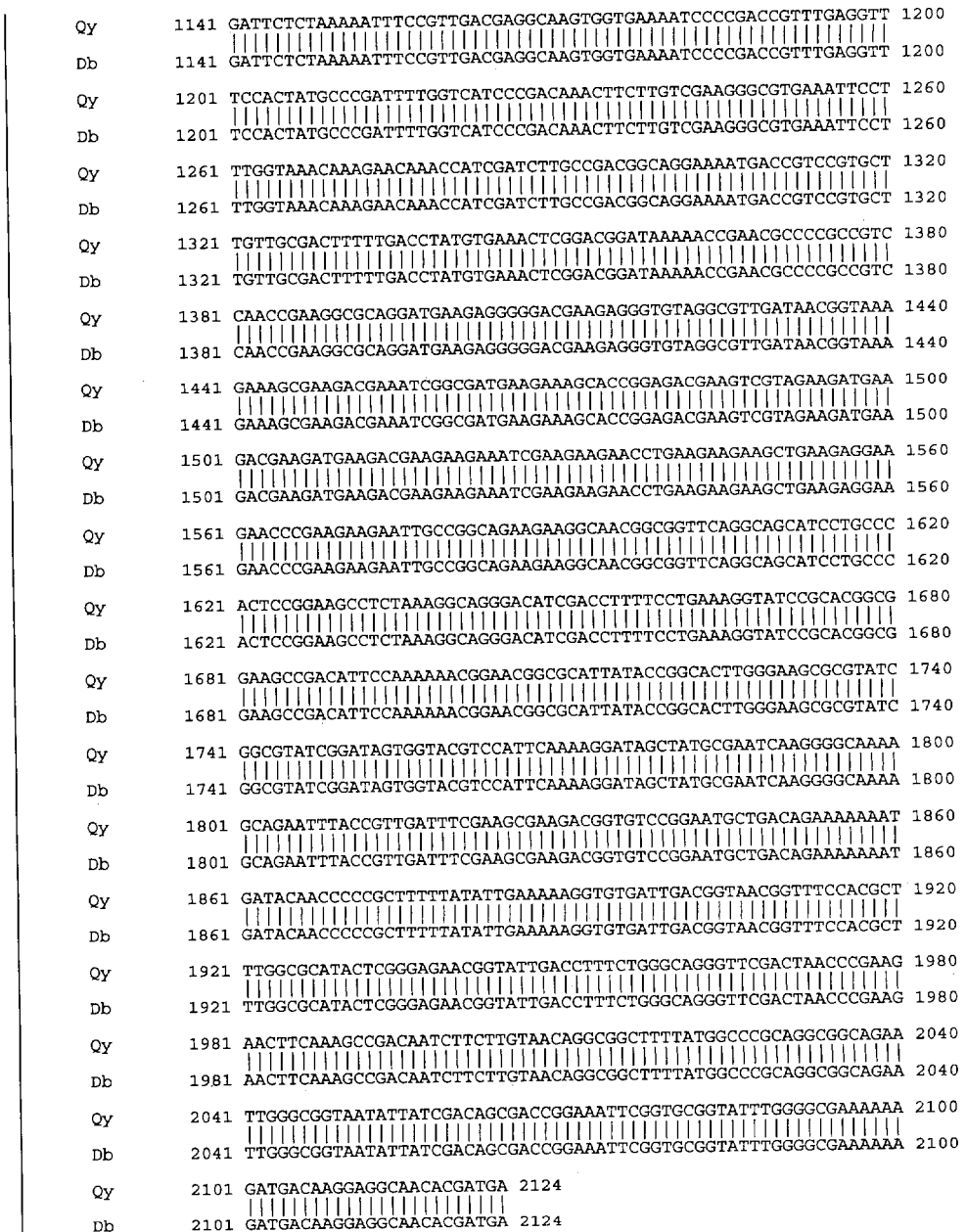
ALIGNMENTS

RESULT 1
US-10-735-098-9
; Sequence 9, Application US/10735098
; Publication No. US20040131634A1
; GENERAL INFORMATION:
; APPLICANT: Pettersson-Fernholm, Annika Margareta
; APPLICANT: Tommassen, Johannes Petrus Maria
; TITLE OF INVENTION: Neisseria Lactoferrin Binding Protein
; FILE REFERENCE: B45106C1
; CURRENT APPLICATION NUMBER: US/10/735,098
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: 09/485,760
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: PCT/EP98/05117
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: GB 9717423.9
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: GB 9805544.8
; PRIOR FILING DATE: 1998-02-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2124
; TYPE: DNA
; ORGANISM: Neisseria meningitidis strain 881607
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2121)
US-10-735-098-9

Query Match 100.0%; Score 2124; DB 17; Length 2124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db 1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60



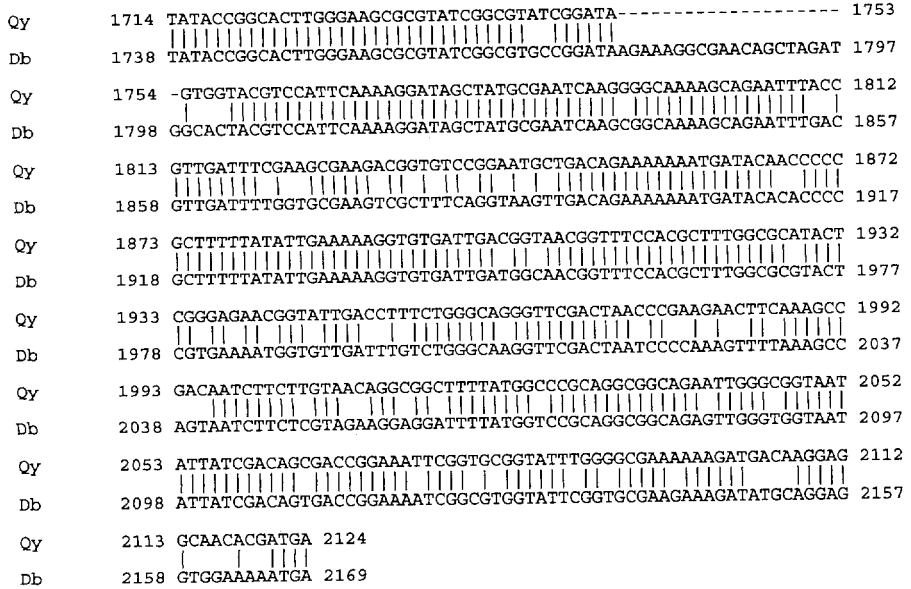


RESULT 2
US-10-735-098-3
; Sequence 3, Application US/10735098
; Publication No. US20040131634A1

; GENERAL INFORMATION:
; APPLICANT: Pettersson-Fernholm, Annika Margareta
; APPLICANT: Tommassen, Johannes Petrus Maria
; TITLE OF INVENTION: Neisseria Lactoferrin Binding Protein
; FILE REFERENCE: B45106C1
; CURRENT APPLICATION NUMBER: US/10/735,098
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: 09/485,760
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: PCT/EP98/05117
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: GB 9717423.9
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: GB 9805544.8
; PRIOR FILING DATE: 1998-02-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Neisseria meningitidis strain M981
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2166)
US-10-735-098-3

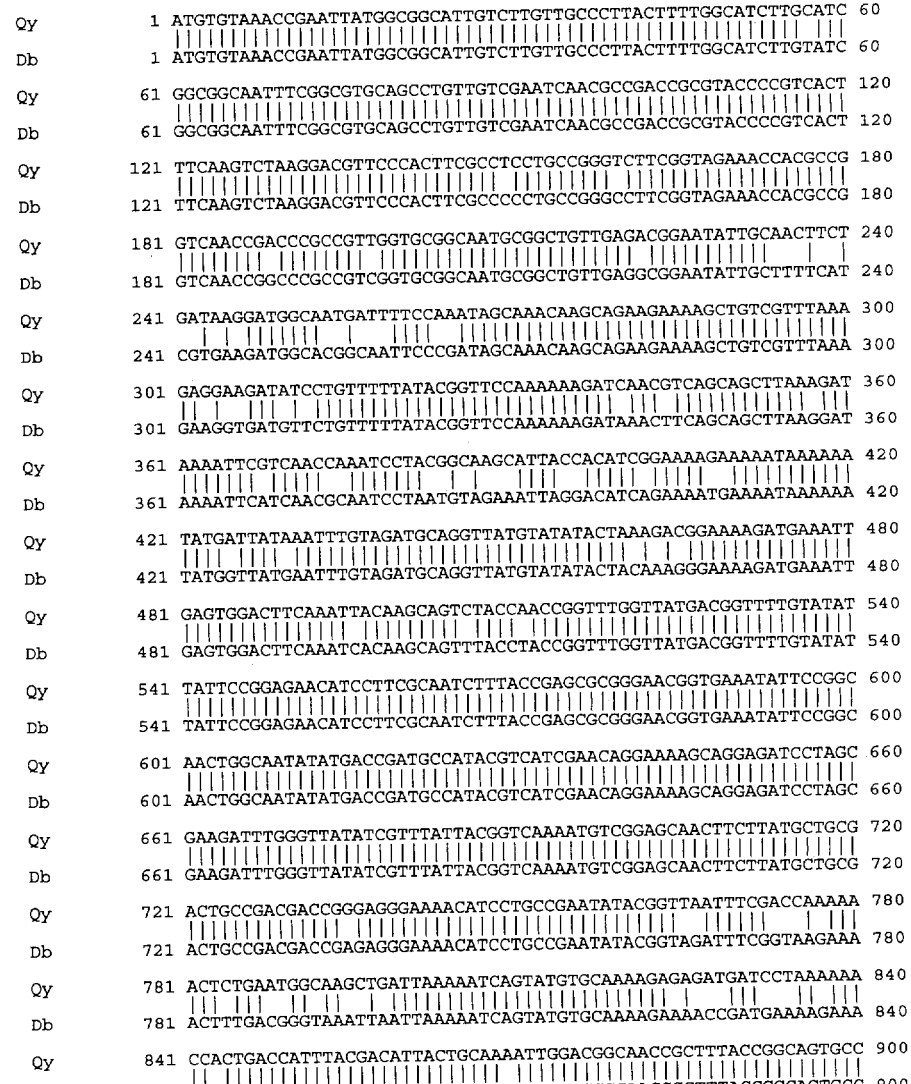
Query Match 73.6%; Score 1562.4; DB 17; Length 2169;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1835; Conservative 0; Mismatches 286; Indels 51; Gaps 5;

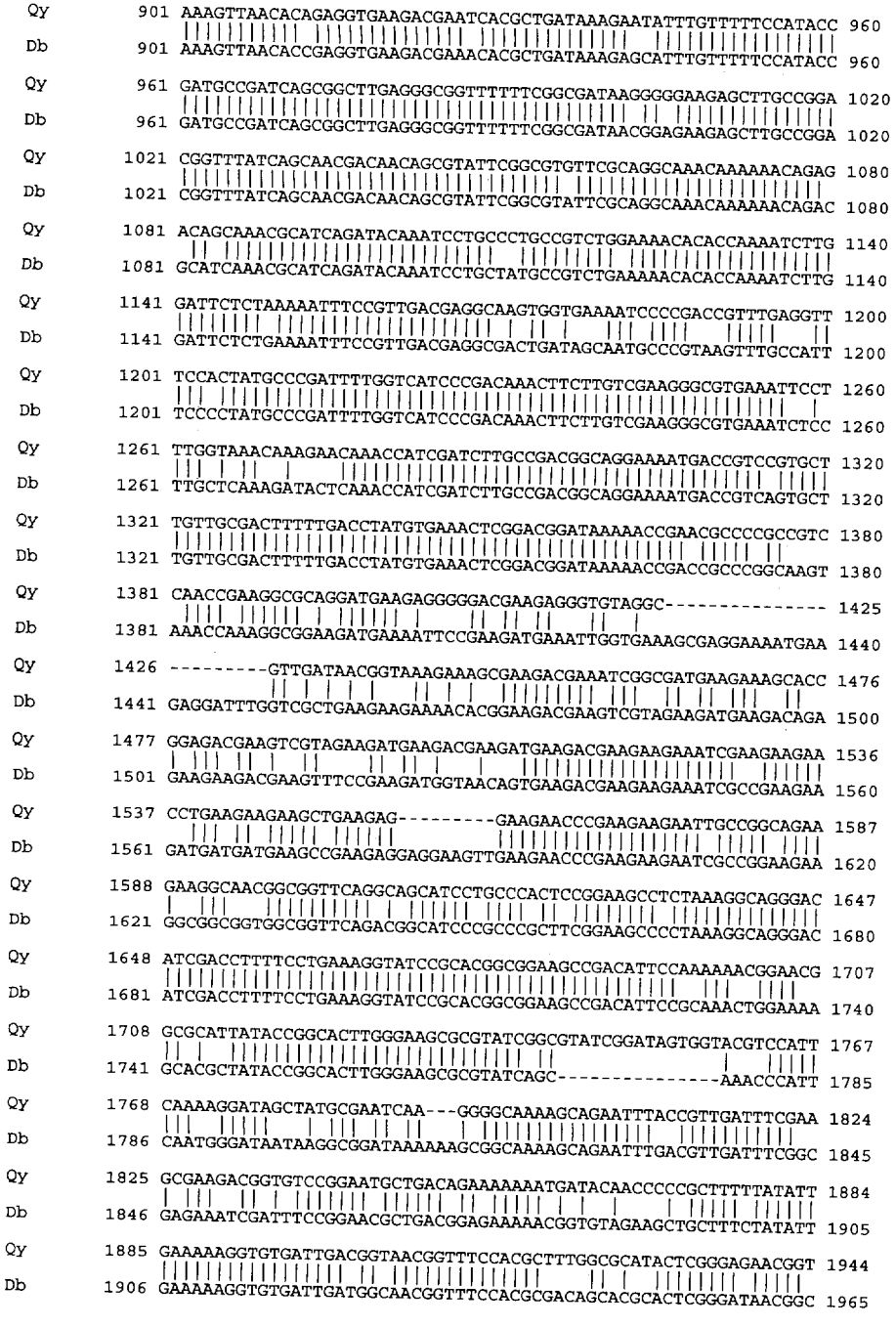
```
Qy    1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
Db    1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60

Qy   61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120
Db   61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCGTACCCCGTCACT 120

Qy  121 TTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180
Db  121 TTCAAGTCTAAGGACGTTCCCACTTCGCCCCCTGCCGGGTCTTCGGTAGAAACCACGCCG 180

Qy  181 GTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAATATTGCAACTTCT 240
Db  181 GTCAACCAGCCCGCCGTCGGTGCGGCAATGCGGCTGTTGAGACGGAATACTGCTTTTCAT 240

Qy  241 GATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300
Db  241 CGTGAAGATGGCACGGCAATTCCCGATAGCAAACAAGCAGAAGAAAAGCTGTCGTTTAAA 300

Qy  301 GAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTAAAGAT 360
Db  301 GAAGGTGATGTTCTGTTTTTATACGGTTCAAAAGAAAATAAACTTCAACAACTTAAAAGC 360

Qy  361 AAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAGAAAAATAAAAAA 420
Db  361 GAAATTCATAAACGTAATCCTGAGGCAAGCATTACCACATCGGAAAATGAAAATAAAAAA 420

Qy  421 TATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGAAAAGATGAAATT 480
Db  421 TATAATTATCGGTTTGTCAGTGCCGGTTATGTGTTTACTAAAAACGGAAAAGATGAAATT 480

Qy  481 GAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGACGGTTTTGTATAT 540
Db  481 GAGAAAACATCGGATGAAAAGCAGTTTTCTAATCGTTTAGGCTATGACGGTTTTGTATAT 540

Qy  541 TATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600
Db  541 TATCTCGGAGAACATCCTTCCCAATCTTTACCGAGCGCGGGAACGGTGAAATATTCCGGC 600

Qy  601 AACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCAGGAGATCCTAGC 660
Db  601 AACTGGCAATATATGACCGATGCCATACGTCATCGGAGAGGTAA---GGGGGTTTCCAGT 657

Qy  661 GAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACTTCTTATGCTGCG 720
Db  658 GTGGATTTGGGTTATACCACATATTATGGTAATGAAATTGGGGCAGCTTCTTATGAGGCT 717

Qy  721 ACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAATTTCGACCAAAAA 780
Db  718 AGGGATGCCGATGGCCGGGAAAAACATCCTGCCGAATATACGGTTAATTTCGACAAAAAA 777

Qy  781 ACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAA 840
Db  778 AACCTGGAAGGTAAGTTGATTAAAAATCAGTATGTGCAAAAGAGAGATGATCCTAAAAAT 837

Qy  841 CCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGCTTTACCGGCAGTGCC 900
Db  838 CCACTGACCATTTACAACATTACCGCAACATTGGACGGCAACCGCTTTACCGGCAGTGCC 897

Qy  901 AAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATATTTGTTTTTCCATACC 960
Db  898 AAAGTTAGCACCGAGGTGAAGACGCAACACGCTGATAAAGAATATTTGTTTTTCCATACC 957

Qy  961 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAAGGGGGAAGAGCTTGCCGGA 1020
Db  958 GATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAACGGAGAAGAGCTTGCCGGG 1017

Qy 1021 CGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1080
Db 1018 CGGTTTATCAGTAACGACAACAGCGTATTCGGCGTGTTCGCAGGCAAACAAAAAACAGAG 1077

Qy 1081 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1140
Db 1078 ACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAACACACCAAAATCTTG 1137

Qy 1141 GATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTT 1200
Db 1138 GATTCTCTAAAAATTTCCGTTGACGAGGCGACTGATGACCATGCCCGTAAGTTTGCCATT 1197

Qy 1201 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1260
Db 1198 TCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCT 1257

Qy 1261 TTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCT 1320
Db 1258 TTGGTTAGCCAAGAGAAAACCATCGAGCTTGCCGACGGCAGGAAAATGACCATCCGTGCT 1317

Qy 1321 TGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACCGAACGCCCCGCCGTC 1380
Db 1318 TGTTGCGATTTTCTGACCTATGTGAAACTCGGACGGATAAAAACCGACCGCCCCGCCGTC 1377

Qy 1381 CAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAA 1440
Db 1378 AAACCGAAGGCGCAGGATGAAGAGGATTCGGACATTGATAATGGCGAAGAAAGCGAAGAC 1437

Qy 1441 GAAA---GCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGA---------- 1481
Db 1438 GAAATTTCCGAAGATGATAACGGCGAAGATGAAGTCACCGAAGAAGAGGAAGCTGAAGAA 1497

Qy 1482 --CGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAA------GAA 1533
Db 1498 ACCGAAGAAGAAACTGATGAAGACGAAGAGGAAGAACCCGAAGAAACTGAAGAAACTGAA 1557

Qy 1534 GAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGC 1593
Db 1558 GAAACTGAAGAAACTGAAGAAACTGAAGAAACTGAAGAAAAATCGCCGACAGAAGAAGGC 1617

Qy 1594 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1653
Db 1618 AACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCTCTAAAGGCAGGGACATCGAC 1677

Qy 1654 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCAAAAAACGGAACGGCGCAT 1713
Db 1678 CTTTTCCTGAAAGGTATCCGCACGGCGGAAGCCGACATTCCGCAAATTGGAAAAGCACGC 1737
```

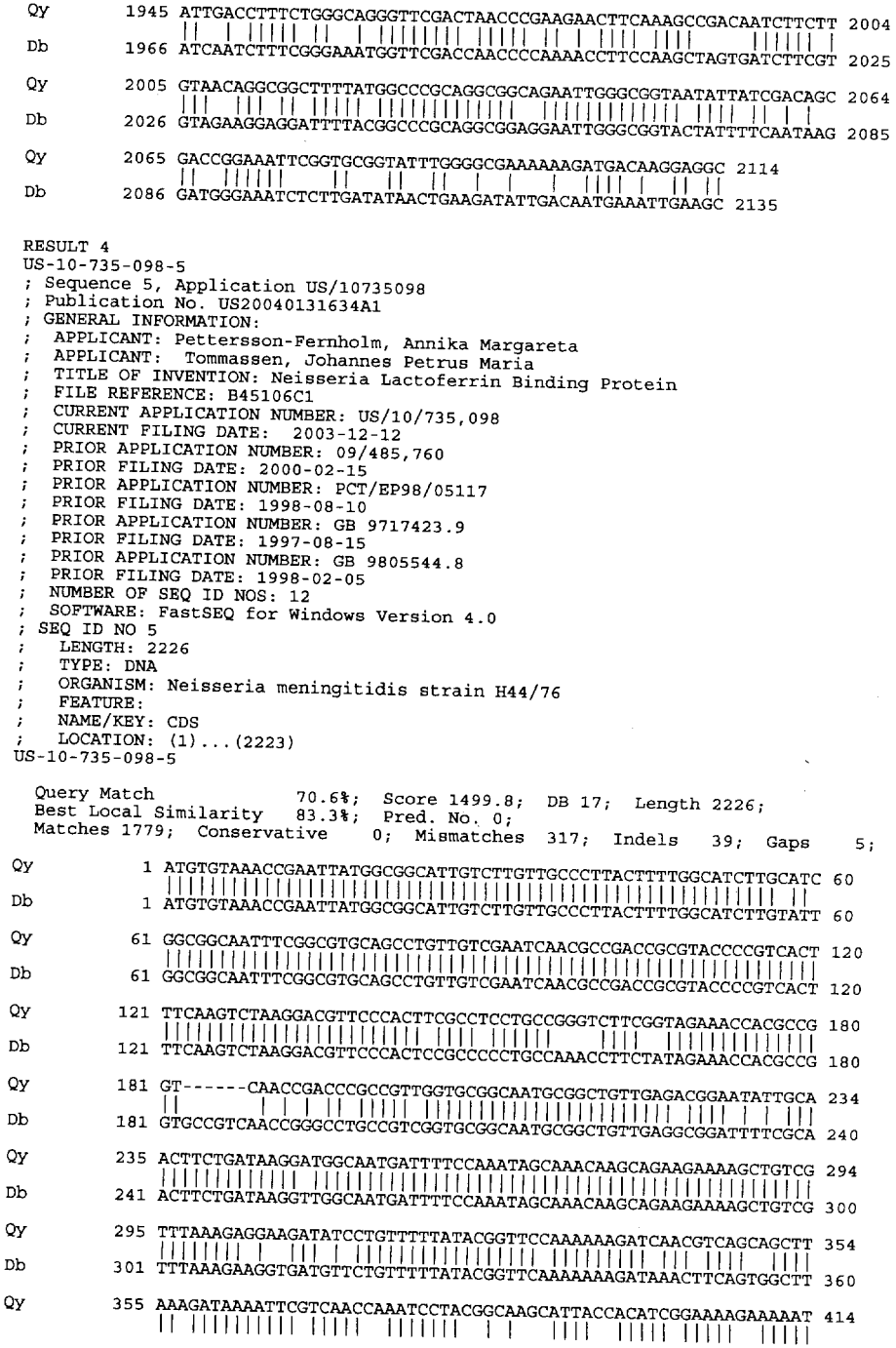


RESULT 3
US-10-282-122A-29815
; Sequence 29815, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29815
; LENGTH: 2226
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-282-122A-29815

Query Match 71.4%; Score 1516.8; DB 13; Length 2226;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 307; Indels 51; Gaps 4;

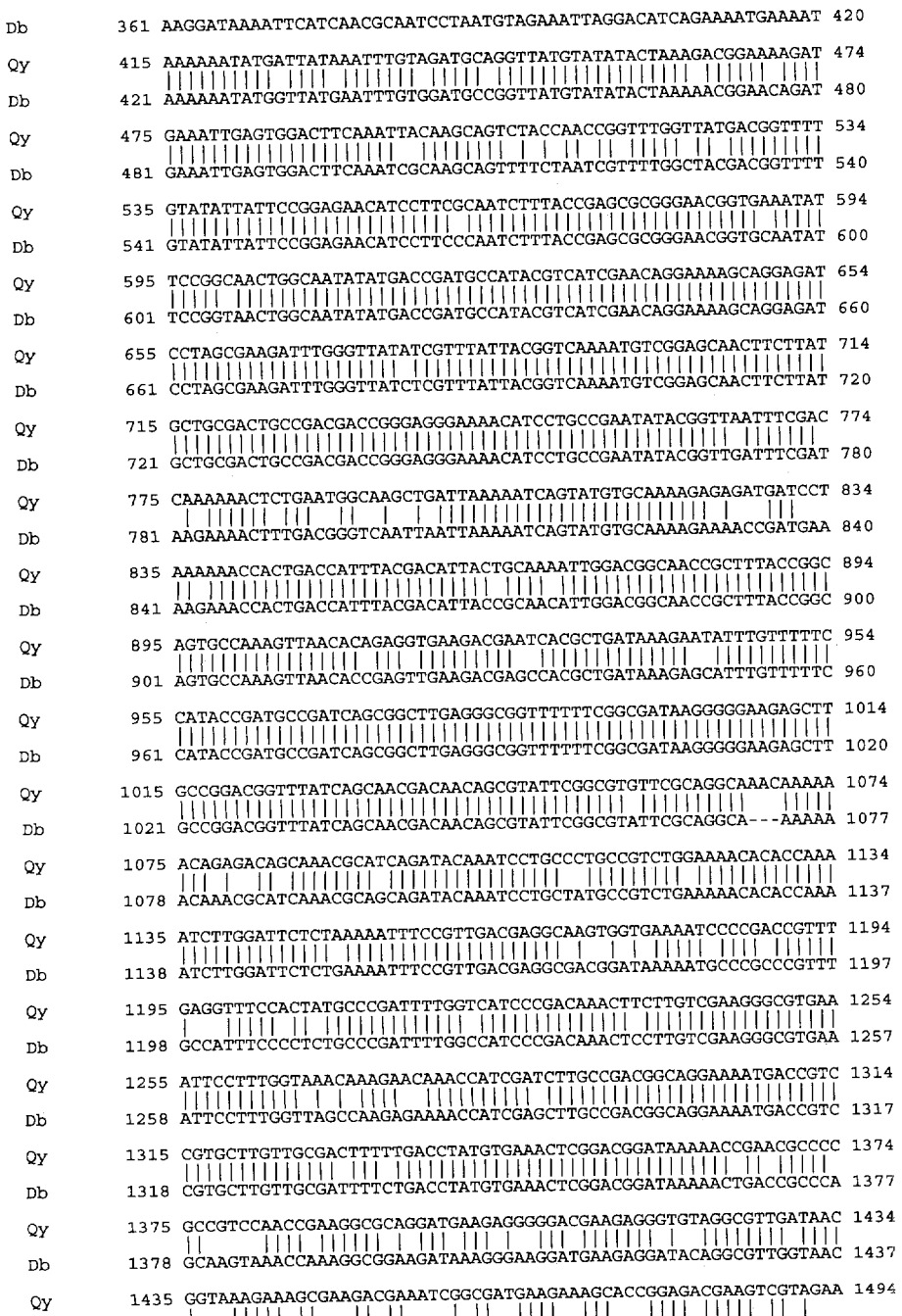


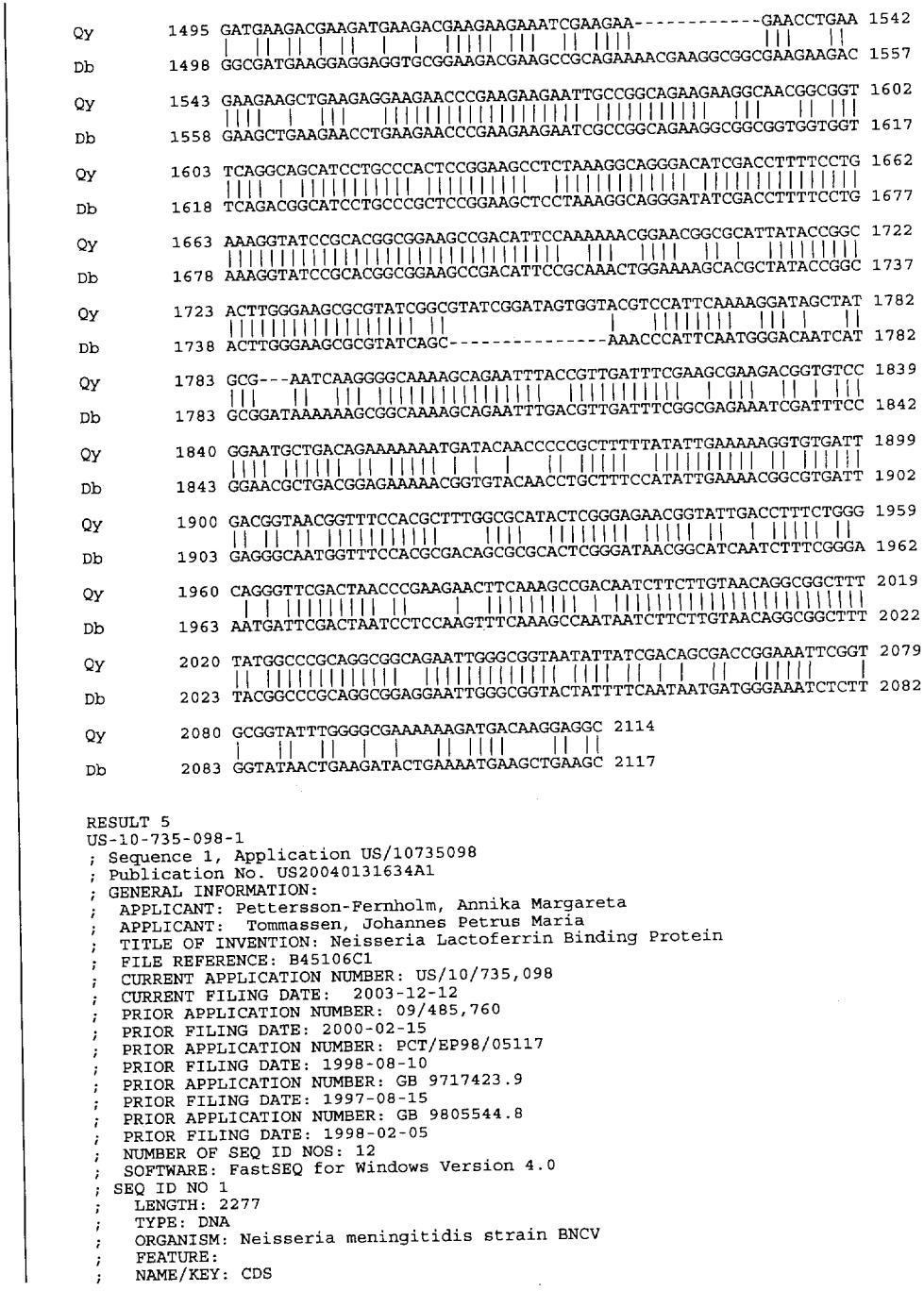




RESULT 4
US-10-735-098-5
; Sequence 5, Application US/10735098
; Publication No. US20040131634A1
; GENERAL INFORMATION:
; APPLICANT: Pettersson-Fernholm, Annika Margareta
; APPLICANT: Tommassen, Johannes Petrus Maria
; TITLE OF INVENTION: Neisseria Lactoferrin Binding Protein
; FILE REFERENCE: B45106C1
; CURRENT APPLICATION NUMBER: US/10/735,098
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: 09/485,760
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: PCT/EP98/05117
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: GB 9717423.9
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: GB 9805544.8
; PRIOR FILING DATE: 1998-02-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2226
; TYPE: DNA
; ORGANISM: Neisseria meningitidis strain H44/76
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2223)
US-10-735-098-5

Query Match 70.6%; Score 1499.8; DB 17; Length 2226;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1779; Conservative 0; Mismatches 317; Indels 39; Gaps 5;

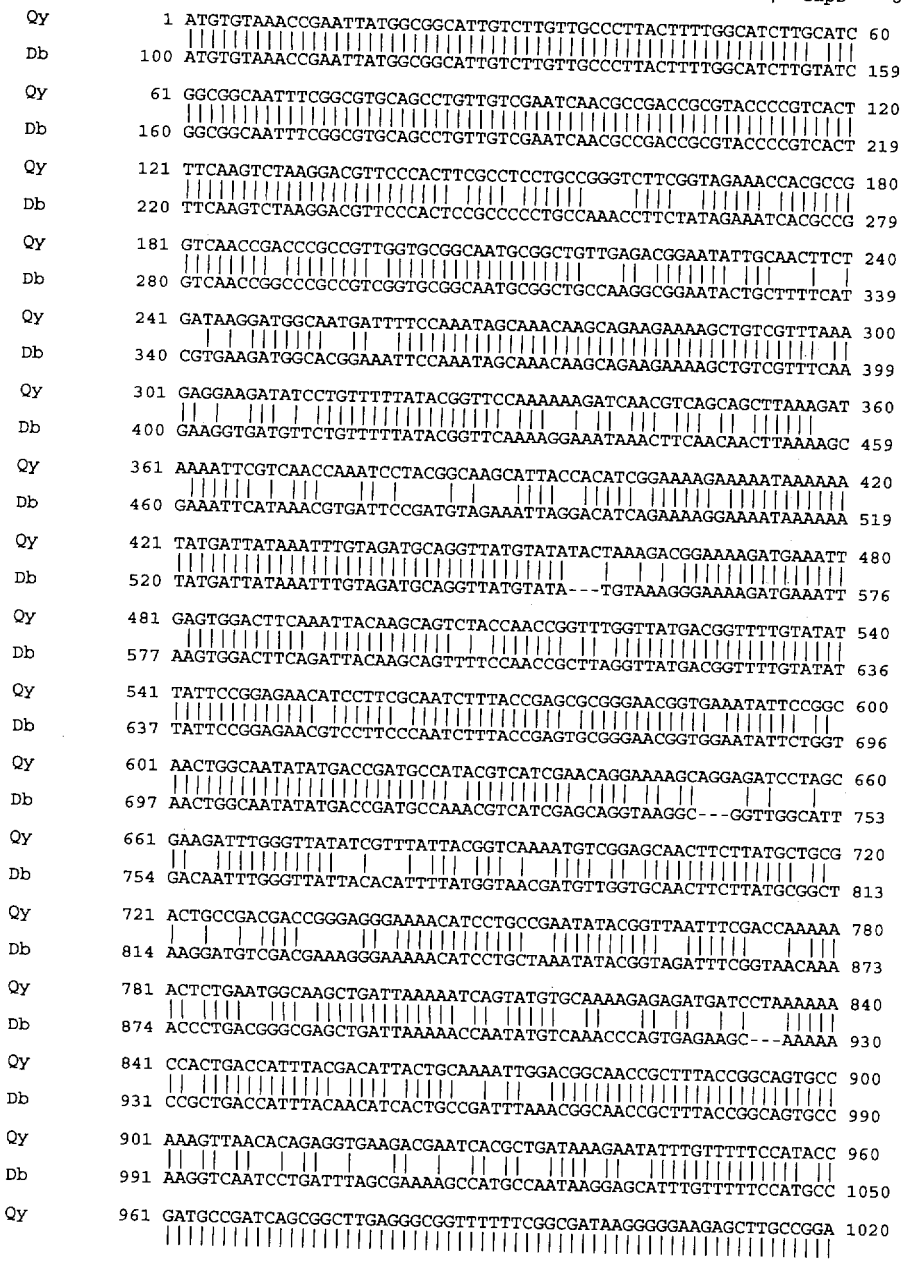


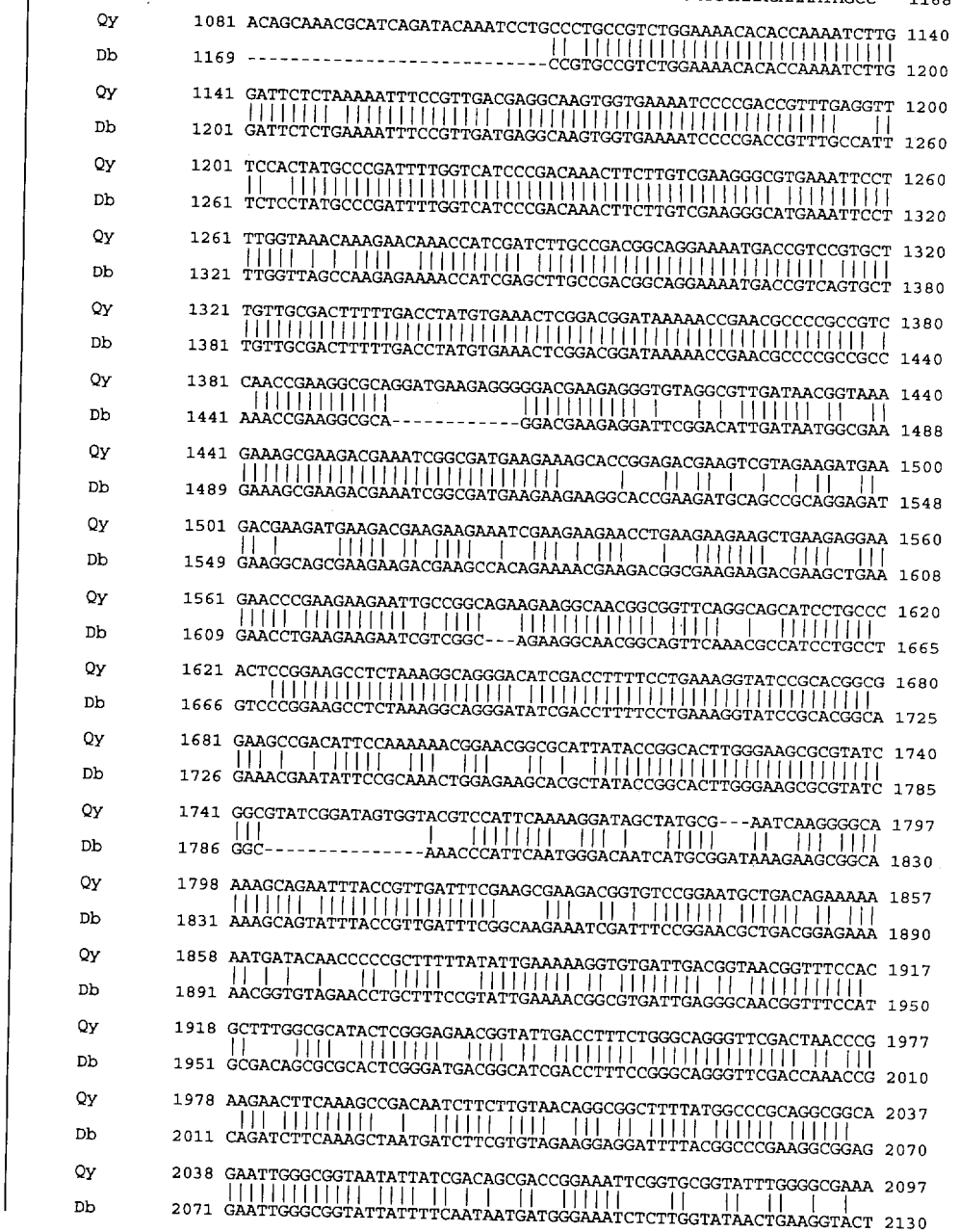


RESULT 5
US-10-735-098-1
; Sequence 1, Application US/10735098
; Publication No. US20040131634A1
; GENERAL INFORMATION:
; APPLICANT: Pettersson-Fernholm, Annika Margareta
; APPLICANT: Tommassen, Johannes Petrus Maria
; TITLE OF INVENTION: Neisseria Lactoferrin Binding Protein
; FILE REFERENCE: B45106C1
; CURRENT APPLICATION NUMBER: US/10/735,098
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: 09/485,760
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: PCT/EP98/05117
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: GB 9717423.9
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: GB 9805544.8
; PRIOR FILING DATE: 1998-02-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2277
; TYPE: DNA
; ORGANISM: Neisseria meningitidis strain BNCV
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (100)...(2274)
US-10-735-098-1

Query Match 61.0%; Score 1296.2; DB 17; Length 2277;
Best Local Similarity 79.0%; Pred. No. 0;
Matches 1672; Conservative 0; Mismatches 373; Indels 72; Gaps 8;



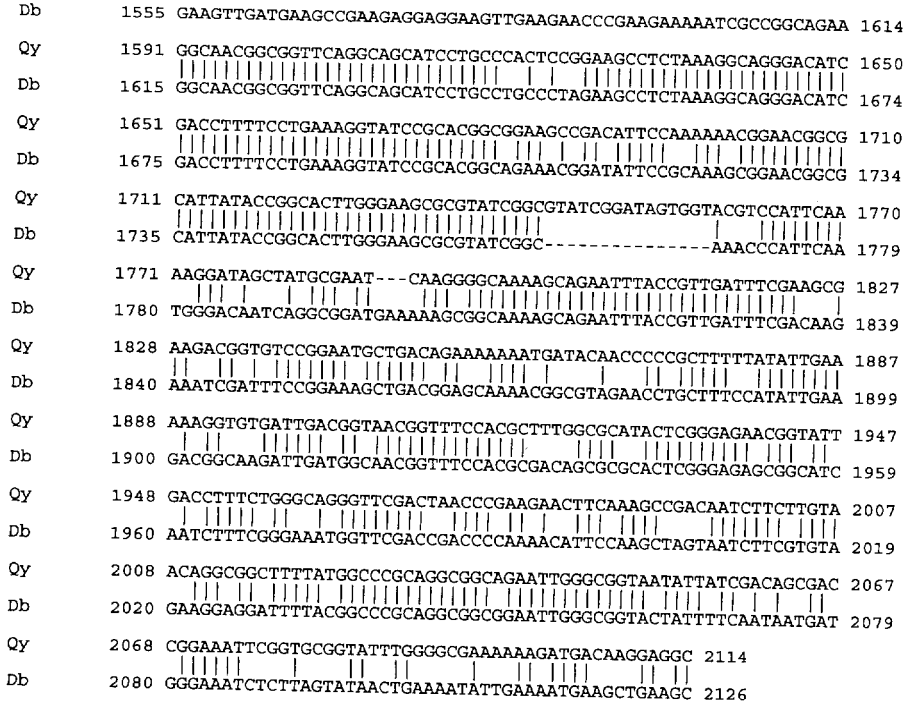


```
Qy       2098 AAAGATGACAAGGAGGC 2114
              || || |    || ||
Db       2131 GAAAATAAAGTTGAAGC 2147
```

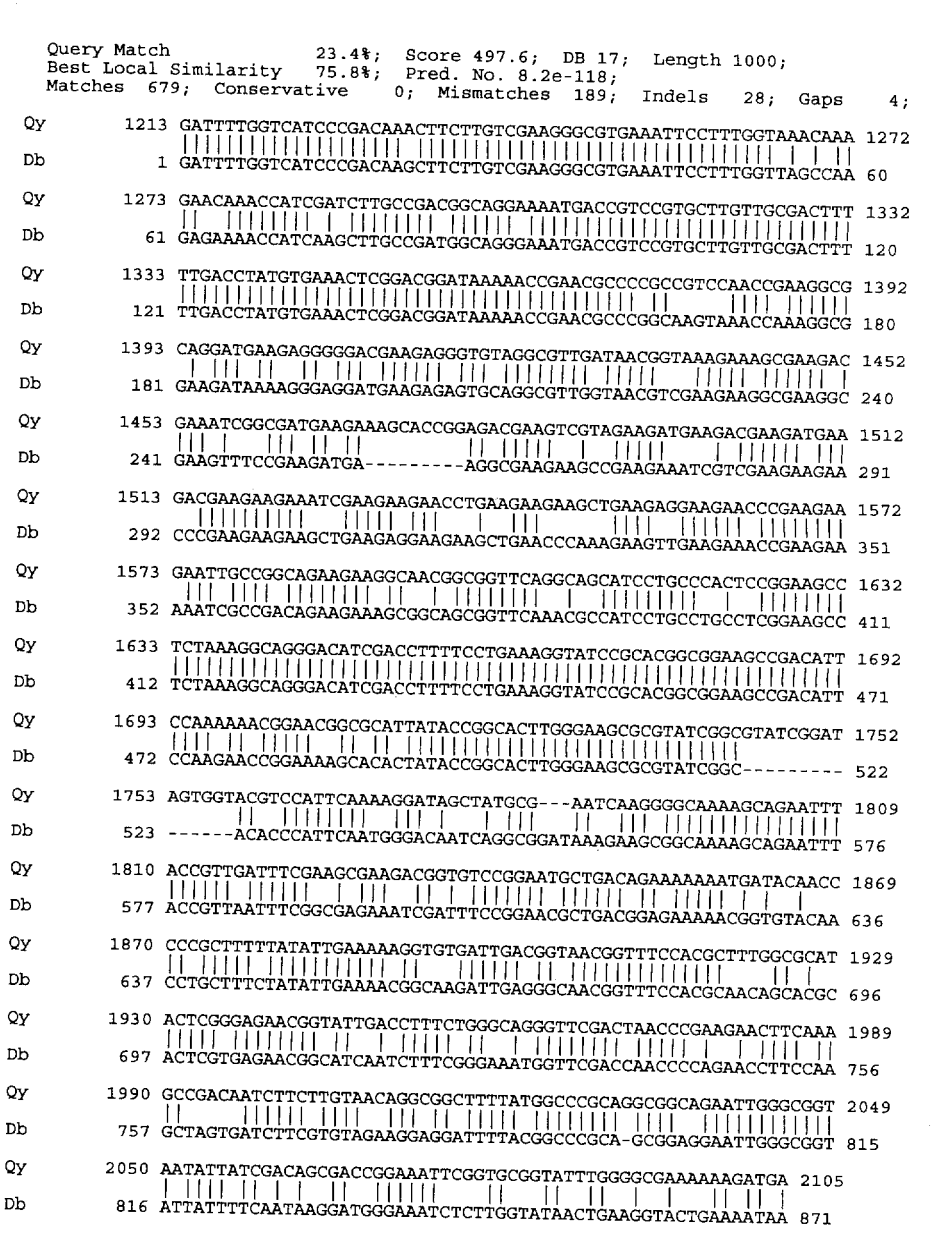
RESULT 6
US-10-735-098-7
; Sequence 7, Application US/10735098
; Publication No. US20040131634A1
; GENERAL INFORMATION:
; APPLICANT: Pettersson-Fernholm, Annika Margareta
; APPLICANT: Tommassen, Johannes Petrus Maria
; TITLE OF INVENTION: Neisseria Lactoferrin Binding Protein
; FILE REFERENCE: B45106C1
; CURRENT APPLICATION NUMBER: US/10/735,098
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: 09/485,760
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: PCT/EP98/05117
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: GB 9717423.9
; PRIOR FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: GB 9805544.8
; PRIOR FILING DATE: 1998-02-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Neisseria meningitidis strain M990
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2259)
US-10-735-098-7

Query Match 60.8%; Score 1291.8; DB 17; Length 2262;
Best Local Similarity 78.1%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 417; Indels 54; Gaps 8;

```
Qy          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTGGCATCTTGCATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||
Db          1 ATGTGTAAACCGAATTATGGCGGCATTGTCTTGTTGCCCTTACTTTTAGCATCTTGTATC 60

Qy         61 GGCGGCAATTTCGGCGTGCAGCCTGTTGTCGAATCAACGCCGACCGCG------------ 108
              |||||||||||||||| |||||||||||||||||||||||||||||||
Db         61 GGCGGCAATTTCGGCGTACAGCCTGTTGTCGAATCAACGCCGACCGCGCCAACTCTGTCA 120

Qy        109 TACCCCGTCACTTTCAAGTCTAAGGACGTTCCCACTTCGCCTCCTGCCGGGTCTTCGGTA 168
              | ||    ||| |||  ||  |||   ||  || |  |||||||| | ||||||||
Db        121 GATTCCAAATCTTCCAATCCTGCGGATAAGCCTGCTCCAGCTCCTGCCGAGCCTTCGGTA 180

Qy        169 GAAACCACGCCGGTCAACCGACCCGCCGTTGGTGCGGCAATGCGGCTGTTGAGACGGAAT 228
              |||| |||||||||||| || |||||||| ||||||||||||||||||    || ||||||
Db        181 GAAATCACGCCGGTCAAGCGGCCCGCCGTCGGTGCGGCAATGCGGCTGCCAAGGCGGAAT 240

Qy        229 ATTGCAACTTCTGATAAGGATGGCAATGATTTTCCAAATAGCAAACAAGCAGAAGAAAAG 288
              || ||||||| ||||||  |||| |||||  |||| ||||| || || ||||| ||  |
Db        241 ATCGCAACTTTTGATAAAAATGGTAATGAAATTCCCAATAGTAAGCAGGCAGAGGAGTAT 300

Qy        289 CTGTCGTTTAAAGAGGAAGATATCCTGTTTTTATACGGTTCCAAAAAAGATCAACGTCAG 348
              ||| || | |||||| | |||||||||||||||| ||||| |     ||||| ||   | |
Db        301 CTGCCGCTCAAAGAGAAGGATATCCTGTTTTTAGACGGTACGCCGAAAGAACAGGCTGAC 360

Qy        349 CAGCTTAAAGATAAAATTCGTCAACCAAATCCTACGGCAAGCATTACCACATCGGAAAAG 408
              | |||||| |  ||||         ||  |||||| |||    ||  ||| || ||
Db        361 AAACTTAAAAAGGAAATCAACGGACGGCATCCTAATGCACCAATCTACACGTCCGATTTA 420

Qy        409 AAAAATAAAAAATATGATTATAAATTTGTAGATGCAGGTTATGTATATACTAAAGACGGA 468
              ||| || |   ||| | ||||||| |||    || || ||||| ||||||| | | |||
Db        421 AAAGATGATGCGTATCAATATAAATATGTCCGGGCCGGATATGTTTATACTAGATATGGA 480

Qy        469 AAAGATGAAATTGAGTGGACTTCAAATTACAAGCAGTCTACCAACCGGTTTGGTTATGAC 528
              |||| |||||  |||| || || |  |||       ||||||    || | | ||||
Db        481 ACAGATGAAATCGAACAGAACTCAGGCGGTAAGCGGGTTACCCACCGCTTAGGTTATGAC 540

Qy        529 GGTTTTGTATATTATTCCGGAGAACATCCTTCGCAATCTTTACCGAGCGCGGGAACGGTG 588
              |||||||||||||||||||||||||||||||| |||||| |||||||||||||| ||||||||||||
Db        541 GGTTTTGTATATTATTCCGGAGAACGTCCTTCCCAATCTTTACCGAGTGCGGGAACGGTG 600

Qy        589 AAATATTCCGGCAACTGGCAATATATGACCGATGCCATACGTCATCGAACAGGAAAAGCA 648
              |||||||  || |||||||||||||||||||||||||| |||||||||| ||||  | ||
Db        601 GAATATTCTGGTAACTGGCAATATATGACCGATGCCAAACGTCATCGAGCAGGTCAGGC- 659

Qy        649 GGAGATCCTAGCGAAGATTTGGGTTATATCGTTTATTACGGTCAAAATGTCGGAGCAACT 708
              | | |  |  ||  |||||||||||||||   | ||| ||| |  |||| || ||||||
Db        660 --GGTTGGCATTGACAATTTGGGTTATATCACATTTTATGGTAACGATGTTGGTGCAACT 717

Qy        709 TCTTATGCTGCGACTGCCGACGACCGGGAGGGAAAACATCCTGCCGAATATACGGTTAAT 768
              |||||||| || |  |  | ||||    || ||| |||||||| ||||||||||  ||
Db        718 TCTTATGCGGCTAAGGATGTCGACGAAAGGGAAAAGCATCCTGCCAAATATACGGTTGAT 777

Qy        769 TTCGACCAAAAAACTCTGAATGGCAAGCTGATTAAAAATCAGTATGTGC---AAAAGAGA 825
              || ||  | |||||  |||||||||||||||||||||||||||||||  | || | |
Db        778 TTTGATAACAAAACCATGAATGGCAAGCTGATTAAAAATCAGTATGTGCGAAATAAAAAA 837

Qy        826 GATGATCCTAAAAAACCACTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGC 885
              ||||| || |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        838 GATGAACCCAAAAAACCGCTGACCATTTACGACATTACTGCAAAATTGGACGGCAACCGC 897

Qy        886 TTTACCGGCAGTGCCAAAGTTAACACAGAGGTGAAGACGAATCACGCTGATAAAGAATAT 945
              |||||||||||||||| || ||  | ||  |   ||  || |  || | ||| ||  |
Db        898 TTTACCGGCAGTGCCAAGGTCAATCCTGATTTAGCGAAAAACCTTGCCGGTAATGAGCGT 957

Qy        946 TTGTTTTTCCATACCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTTCGGCGATAAGGGG 1005
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||| ||
Db        958 TTGTTTTTCCATGCCGATGCCGATCAGCGGCTTGAGGGCGGTTTTTTCGGCGATAACGGA 1017

Qy       1006 GAAGAGCTTGCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTGTTCGCAGGC 1065
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1018 GAAGAGCTTGCCGGACGGTTTATCAGCAACGACAACAGCGTATTCGGCGTATTCGCAGGC 1077

Qy       1066 AAACAAAAAAACAGAGACAGCAAACGCATCAGATACAAATCCTGCCCTGCCGTCTGGAAAA 1125
              |  |||||||||||||||||||||| |||||||||| |||||||||||||||||||||||
Db       1078 A---AAAAAACAGAGACAGCAAACGCAGCAGATACAAAACCTGCCCTGCCGTCTGGAAAA 1134

Qy       1126 CACACCAAAATCTTGGATTCTCTAAAAATTTCCGTTGACGAGGCAAGTGGTGAAAATCCC 1185
              |||||||||||||||||||||||||||||||||||||||||||||| | || || || ||
Db       1135 CACACCAAAATCTTGGATTCTCTAAAAATTTCCGTTGACGAGGCGACTGATGGCCATGCC 1194

Qy       1186 CGACCGTTTGAGGTTTCCACTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAA 1245
              ||   |||||  |||| |||||||||||||||||||||||||||||||||||||||||||
Db       1195 CGTAAGTTTGCCATTTCCTCTATGCCCGATTTTGGTCATCCCGACAAACTTCTTGTCGAA 1254

Qy       1246 GGGCGTGAAATTCCTTTGGTAAACAAAGAACAAACCATCGATCTTGCCGACGGCAGGAAA 1305
              |||||||||||||||||||| |||||||| |||| | ||||||||||||||||||||
Db       1255 GGGCGTGAAATTCCTTTGGTAAACGAAGAACAAATCATCAAGCTTGCCGACGGCAGGAAA 1314

Qy       1306 ATGACCGTCCGTGCTTGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACC 1365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 ATGACCGTCCGTGCTTGTTGCGACTTTTTGACCTATGTGAAACTCGGACGGATAAAAACC 1374

Qy       1366 GAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGC 1425
              || |||| ||  |||| ||||||| | ||| || ||| || ||| |||||||||| ||||
Db       1375 GATCGCCCGGCAAGTAAACCAAAGGCGGAAGATAAAGGGGAGGATGAAGAGGGTGCAGGC 1434

Qy       1426 GTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAA 1485
              |||||||||  |||||||||||||||||||  || || || ||| || || ||| ||
Db       1435 GTTGATAACGACGAAGAAAGCGAAGACGAAGCCGTAGAAGACGAAGGCGGCGAAGAAGAC 1494

Qy       1486 G------TCGTAGAAGATGAAGACGAAGATGAAGA---------CGAAGAAGAAATCGAA 1530
              |      || ||| ||| | | |||||| |||||           |||||||||| ||||
Db       1495 GAAACTTCCGAAGAGGATAATGGCGAAGACGAAGAAGCAACCGCCGAAGAAGAAACCGAA 1554

Qy       1531 GAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAA 1590
              ||||    |||||  ||||  || || |  |||       ||||||    || | | ||||
```



RESULT 7
US-10-343-561-15
; Sequence 15, Application US/10343561
; Publication No. US20040126389A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Dalemans, Wilfried
; APPLICANT: Denoel, Philippe
; APPLICANT: Dequesne, Guy
; APPLICANT: Feron, Christiane
; APPLICANT: Garcon, Nathalie
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Thiry, Georges
; APPLICANT: Thonnard, Joelle
; APPLICANT: Voet, Pierre
; TITLE OF INVENTION: Vaccines Comprising Outer Membrane
; TITLE OF INVENTION: Vesicles from Gram Negative Bacteria
; FILE REFERENCE: B45260
; CURRENT APPLICATION NUMBER: US/10/343,561
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/EP01/08857
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: EP 00956369.3
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: GB 0103170.7
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-343-561-15



RESULT 8
US-10-467-534-80
; Sequence 80, Application US/10467534
; Publication No. US20040131625A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves

```
;  APPLICANT:  Poolman, Jan
;  APPLICANT:  Verlant, Vincent Georges Christian Louis
;  TITLE OF INVENTION: Vaccine Composition
;  FILE REFERENCE: B45261
;  CURRENT APPLICATION NUMBER: US/10/467,534
;  CURRENT FILING DATE:  2003-08-08
;  PRIOR APPLICATION NUMBER: PCT/EP02/01356
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: GB 0103169.9
;  PRIOR FILING DATE: 2001-02-08
;  NUMBER OF SEQ ID NOS: 108
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
;   LENGTH: 3300
;   TYPE: DNA
;   ORGANISM: Neisseria gonorrhoeae
US-10-467-534-80

  Query Match             4.9%;  Score 104.6;  DB 17;  Length 3300;
  Best Local Similarity   74.9%;  Pred. No. 6.6e-16;
  Matches  131;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1931 CTCGGGAGAACGGTATTGACCTTTCTGGGCAGGGTTCGACTAACCCGAAGAACTTCAAAG 1990
             ||||||| ||||| ||  | ||||| ||  | |||||||||| ||  | |  |||||||
Db          1 CTCGGGATAACGGCATCAATCTTTCGGGAAATGGTTCGACTAATCCTCAAAGTTTCAAAG 60

Qy       1991 CCGACAATCTTCTTGTAACAGGCGGCTTTTATGGCCCGCAGGCGGCAGAATTGGGCGGTA 2050
             |||||||||||||||||| |||||||||| |||||||||||||| |||||||||| |
Db         61 CCGACAATCTTCTTGTAACGGGCGGCTTTTACGGCCCGCAGGCGGCGGAATTGGGCGGCA 120

Qy       2051 ATATTATCGACAGCGACCGGAAATTCGGTGCGGTATTTGGGGCGAAAAAAGATGA 2105
              |||| || | |  ||  |||||||    ||  ||  ||  |  |   || ||||
Db        121 CTATTTTCAATAAGGATGGGAAATCTCTTGGTATAACTGAAGATATTGAAAATGA 175


RESULT 9
US-10-087-192-1627/c
; Sequence 1627, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1627
;   LENGTH: 374849
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(374849)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1627

  Query Match             4.6%;  Score 98.4;  DB 13;  Length 374849;
  Best Local Similarity   66.5%;  Pred. No. 4.6e-13;
  Matches  141;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
             | || |||||||  || |||||||  | ||  |  ||  | ||  |||||  | |||||  |
Db     366742 AAGAGGAAGAGGAAGAAGAAGAGGAAGAAGAAGGAGAAGAAGGAGAAGAAGAAGGAGAAGAAG 366683

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
                |   |  |  |||||| |    | ||| ||||  | |||||| |||||  |||||  ||||

Db     366682 GAGAAGAAGGAGAAGAAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 366623

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
              | |||||||  |   | |||||  |  ||||||||||  |||||  ||||||   |||||||
Db     366622 AAGAAGAAGGAGGAGGAGAAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 366563

Qy       1574 AATTGCCGGCAGAAGAAGGCAACGGCGGTTCA 1605
              ||       | |||||||  ||||   | || |
Db     366562 AAGAAGAAGAAGAAGAAAACAACCCTGCTTTA 366531


RESULT 10
US-10-087-192-463/c
; Sequence 463, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 463
;   LENGTH: 31124
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(31124)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-463

  Query Match             4.6%;  Score 97.2;  DB 13;  Length 31124;
  Best Local Similarity   68.2%;  Pred. No. 2.1e-13;
  Matches  135;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
             | || ||||| |  || ||||| |  | ||  |  ||  | |  |||||  |||||  |
Db      18558 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 18499

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
              ||   |  |  |||||| |    | ||||  ||||  | |||||| |||||  ||||

Db      18498 AAGAAGAAGGAGAAGAAGGAGGAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 18439

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
              | |||||||||   |||||||||   ||||||||||  |||||  ||||||   |||||||
Db      18438 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 18379

Qy       1574 AATTGCCGGCAGAAGAAG 1591
              ||       | ||||||||
Db      18378 AAGAAGAAGAAGAAGAAG 18361


RESULT 11
US-10-087-192-49
; Sequence 49, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
```

```
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
;   LENGTH: 37265
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(37265)
;   OTHER INFORMATION: n = A,T,C or G
US-10-087-192-49

  Query Match             4.5%;  Score 96.6;  DB 13;  Length 37265;
  Best Local Similarity   66.7%;  Pred. No. 3.4e-13;
  Matches  138;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
             | || | ||| |  || | ||| |  | ||| |  ||  | ||  |||||  | ||| |
Db      29840 AAGAAGGAGAAGAAGAAGGAGAAGAAGAAGGAGAAGAAGAAGGAGAAGAAGAAGGAGAAG 29899

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
             ||   || || |||||| |   | ||| |||| | |||||| ||||| ||||| ||||
Db      29900 AAGAAGGAGAAGAAGAAGGAGAAGAAGAAGAAGAAGGAGAAGAAGAAGAAGAAGAAGAAG 29959

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
             | |||||||||  |||||||||  |||||||||  |||||  ||||||    |||||||
Db      29960 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGGAGAAGAAG 30019

Qy       1574 AATTGCCGGCAGAAGAAGGCAACGGCG 1600
             ||      | || |||||   | || |
Db      30020 AAGAAGAAGAAGGAGAAGAAGAAGGAG 30046


RESULT 12
US-10-322-281-773/c
; Sequence 773, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc S. Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001000
;  CURRENT APPLICATION NUMBER: US/10/322,281
;  CURRENT FILING DATE:  2002-12-17
;  NUMBER OF SEQ ID NOS: 866
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 773
;   LENGTH: 115223
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(115223)
;   OTHER INFORMATION: n = A,T,C or G
US-10-322-281-773

  Query Match             4.5%;  Score 96;  DB 17;  Length 115223;
  Best Local Similarity   32.8%;  Pred. No. 9.5e-13;
  Matches  150;  Conservative    0;  Mismatches  308;  Indels    0;  Gaps    0;

Qy       1165 GAGGCAAGTGGTGAAAATCCCCGACCGTTTGAGGTTTCCACTATGCCCGATTTTGGTCAT 1224
             |||| |
Db       9462 GAGGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9403

Qy       1225 CCCGACAAACTTCTTGTCGAAGGGCGTGAAATTCCTTTGGTAAACAAAGAACAAACCATC 1284

Db       9402 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9343

Qy       1285 GATCTTGCCGACGGCAGGAAAATGACCGTCCGTGCTTGTTGCGACTTTTTGACCTATGTG 1344

Db       9342 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9283

Qy       1345 AAACTCGGACGGATAAAAACCGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAG 1404
                                                                 |    || | |||
Db       9282 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAGAAGGAGAAGGAGAA 9223

Qy       1405 GGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGAT 1464
              || || | ||| || | ||| | |  | || ||| |  | ||| | |  || ||
Db       9222 GGAGAAGGAGAAGGAGAAGGAGAAGGAGAAGGAGAAGGAGAAGGAGAAGGAGAAGGAGAA 9163

Qy       1465 GAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAA 1524
              | |||| |    |||||  | ||  | |||||  |||| |||||  ||||| |||||||||
Db       9162 GGAGAAGGAGAAGGAGAAGGAGAAGGAGAAGGAGAAGGAGAAGAAGAAGAAGAAGAAGAA 9103

Qy       1525 ATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCA 1584
              |||||||||  ||||||||||  |||||  ||||||  |||||||||       | |
Db       9102 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 9043

Qy       1585 GAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCCAC 1622
              |||||||  ||  | |  |||  | | |  | | | |
Db       9042 GAAGAAGAAAAGAGAGAGACAGAGACCTTGGTCCTCGC 9005


RESULT 13
US-10-027-632-269927/c
; Sequence 269927, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 269927
;   LENGTH: 635
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-269927

  Query Match             4.5%;  Score 95.4;  DB 13;  Length 635;
  Best Local Similarity   66.5%;  Pred. No. 5.9e-14;
  Matches  135;  Conservative    1;  Mismatches   67;  Indels    0;  Gaps    0;

Qy       1388 AGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCG 1447
              |||   |||| ||||| || || | ||| || | ||  |  ||  | |  |||||   |
Db        425 AGGAAGAGGAAGAAGAAGGAGAAGGAGAAGGAGGAGAAGGAGAAGAAGAAGAAGAAGAAG 366

Qy       1448 AAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAG 1507
              |||| |||   |  || ||||||      | ||| ||||  | ||||:| |||||  |||
Db        365 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAACAAG 306

Qy       1508 ATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCG 1567
              | |||||  ||||||||  |||||||||  ||||||||||  |||||  |||||   |
Db        305 AAGAAGAACAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 246

Qy       1568 AAGAAGAATTGCCGGCAGAAGAA 1590
```

```
                ||||||||        | |||||||
Db        245 AAGAAGAAGAAGAAGAAGAAGAA 223


RESULT 14
US-10-027-632-269927/c
; Sequence 269927, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 269927
;   LENGTH: 635
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-269927

  Query Match             4.5%;  Score 95.4;  DB 16;  Length 635;
  Best Local Similarity   66.5%;  Pred. No. 5.9e-14;
  Matches  135;  Conservative    1;  Mismatches   67;  Indels    0;  Gaps    0;

Qy       1388 AGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCG 1447
             |||   |||| ||||| || || | ||| || | ||  |  ||  | |   |||||   |
Db        425 AGGAAGAGGAAGAAGAAGGAGAAGGAGAAGGAGGAGAAGGAGAAGAAGAAGAAGAAGAAG 366

Qy       1448 AAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAG 1507
             |||| |||   |  || ||||||      | ||| |||| | ||||:| ||||| |||
Db        365 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAACAAG 306

Qy       1508 ATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCG 1567
             | |||||  |||||||||   |||||||||   ||||||||||  ||||| ||||||  |
Db        305 AAGAAGAACAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 246

Qy       1568 AAGAAGAATTGCCGGCAGAAGAA 1590
             ||||||||      | |||||||
Db        245 AAGAAGAAGAAGAAGAAGAAGAA 223


RESULT 15
US-09-864-761-19262/c
; Sequence 19262, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19262
;   LENGTH: 305
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC011416.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL =  35
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
;   OTHER INFORMATION: SWISSPROT HIT: P05156, EVALUE 9.40e+00
;   OTHER INFORMATION: NT HIT: U67539.1, EVALUE 7.40e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BF129635.1, EVALUE 1.10e+00
US-09-864-761-19262

  Query Match             4.4%;  Score 93.8;  DB 9;  Length 305;
  Best Local Similarity   65.6%;  Pred. No. 9.8e-14;
  Matches  137;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1389 GGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGA 1448
             || |  |||   |||| | ||  ||||| | | ||  |  ||  | |  |||||   ||
Db        223 GGTGGGGGAGAGAGAGCGAGAGAAAGAGAGAGAAGAAGAAGAAGAAGAAGAAGAAGA 164

Qy       1449 AGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGA 1508
             ||| |||   |  || ||||||      | ||| ||||  | |||||| ||||| |||||
Db        163 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 104

Qy       1509 TGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGA 1568
              ||||| |||||||||   |||||||||   ||||||||||  ||||| ||||||   ||
Db        103 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 44
```

```
Qy       1569 AGAAGAATTGCCGGCAGAAGAAGGCAACG 1597
              ||||||| |||| ||  || |
Db         43 AGAAGAAGAAGAAGAAGAAAAATTTAAAG 15
```

Search completed: August 25, 2004, 14:58:25
Job time : 1076.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 11:49:00 ; Search time 4846.67 Seconds
(without alignments)
13086.780 Million cell updates/sec

Title: US-10-735-098-9
Perfect score: 2124
Sequence: 1 atgtgtaaaccgaattatgg.....acaaggaggcaacacgatga 2124

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	103.4	4.9	531	14	CD915362	CD915362 G550.125O
	2	101	4.8	802	28	BH066160	BH066160 RPCI-24-2
c	3	100.8	4.7	434	14	CA745444	CA745444 wri2s.pk0
c	4	99.4	4.7	289	14	CA716220	CA716220 wdk3c.pk0
	5	99.2	4.7	223	13	CA022136	CA022136 HZ42E07r
	6	99	4.7	395	29	CE760553	CE760553 tigr-gss-
	7	98.8	4.7	264	13	BQ241566	BQ241566 TaE05003E
	8	98.6	4.6	729	29	CE180020	CE180020 tigr-gss-
c	9	98.6	4.6	1096	14	CK206118	CK206118 FGAS01769
	10	97.6	4.6	799	28	BZ201507	BZ201507 CH230-445
	11	97.4	4.6	575	10	BF707975	BF707975 A634_LE A
c	12	97.2	4.6	232	29	CG321416	CG321416 OGYBI53TV
c	13	97.2	4.6	233	12	BJ454017	BJ454017 BJ454017
c	14	97.2	4.6	691	28	BH040136	BH040136 RPCI-24-2
	15	97	4.6	675	28	BZ104011	BZ104011 CH230-227
	16	96.8	4.6	278	13	BQ460618	BQ460618 HA06H01r
	17	96.8	4.6	744	28	BH065422	BH065422 RPCI-24-2
c	18	96.8	4.6	774	28	AZ658872	AZ658872 1M0535G22
	19	96.6	4.5	477	14	CD311846	CD311846 StrPu691.
	20	96.6	4.5	960	28	CC143364	CC143364 NDL.32A13
c	21	96.4	4.5	263	13	BU005139	BU005139 QGG7D20.y
	22	96.2	4.5	719	28	BH360392	BH360392 CH230-107
c	23	95.8	4.5	311	14	CA699402	CA699402 wlk8.pk00
	24	95.8	4.5	980	13	BU944408	BU944408 AGENCOURT
	25	95.6	4.5	332	14	CA566759	CA566759 K0405E07-
c	26	95.6	4.5	605	28	BZ222147	BZ222147 CH230-361
	27	95.6	4.5	617	14	CA743799	CA743799 wrils.pk0
	28	95.4	4.5	205	13	BQ241058	BQ241058 TaE05009F
c	29	95.4	4.5	546	28	AZ600777	AZ600777 1M0418D18
	30	94.8	4.5	341	29	CE510654	CE510654 tigr-gss-
	31	94.6	4.5	924	13	BQ925195	BQ925195 AGENCOURT
	32	94.2	4.4	340	13	BQ462203	BQ462203 HD02A20r
c	33	94	4.4	483	28	AZ060017	AZ060017 RPCI-23-4
	34	94	4.4	527	28	AZ279446	AZ279446 RPCI-23-1
	35	94	4.4	596	29	CC660549	CC660549 OGWIL63TV
c	36	93.8	4.4	613	29	CE357285	CE357285 tigr-gss-
c	37	93.8	4.4	880	28	AZ529191	AZ529191 ENTBV68TR
	38	93.8	4.4	953	13	BQ939407	BQ939407 AGENCOURT
	39	93.6	4.4	293	13	BQ459561	BQ459561 HA08M06r
	40	93.6	4.4	295	13	BU976670	BU976670 HA10B15r
	41	93.6	4.4	665	13	BU432980	BU432980 603221001
	42	93	4.4	342	11	AY383662	AY383662 Rattus no
	43	93	4.4	537	14	CF380795	CF380795 lac78f12.
	44	92.8	4.4	358	28	AZ582992	AZ582992 1M0376K12
	45	92.8	4.4	495	14	CD663452	CD663452 UCRHV18_0

ALIGNMENTS

RESULT 1
CD915362
LOCUS CD915362 531 bp mRNA linear EST 14-JUL-2003
DEFINITION G550.125O01F010713 G550 Triticum aestivum cDNA clone G550125O01, mRNA sequence.
ACCESSION CD915362
VERSION CD915362.1 GI:32689686
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 531)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (http://www.genoplante.com and http://genoplante-info.infobiogen.fr).
FEATURES Location/Qualifiers
source 1. .531

/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G550125O01"
/tissue_type="grain (550 degrees per day after pollination)"
/clone_lib="G550"

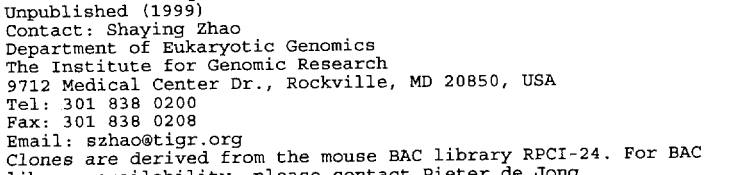
ORIGIN

Query Match 4.9%; Score 103.4; DB 14; Length 531;
Best Local Similarity 67.3%; Pred. No. 2.7e-10;
Matches 146; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

RESULT 2
BH066160
LOCUS BH066160 802 bp DNA linear GSS 18-JUL-2001
DEFINITION RPCI-24-285J5.TJ RPCI-24 Mus musculus genomic clone RPCI-24-285J5, genomic survey sequence.
ACCESSION BH066160
VERSION BH066160.1 GI:14884017
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 802)
AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M., Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E., Russell,D., de Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-24
JOURNAL Unpublished (1999)
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 285 row: J column: 5
Seq primer: SP6
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .802
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-285J5"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA."

ORIGIN

Query Match 4.8%; Score 101; DB 28; Length 802;
Best Local Similarity 65.1%; Pred. No. 8.3e-10;
Matches 149; Conservative 0; Mismatches 80; Indels 0; Gaps 0;



RESULT 3
CA745444/c
LOCUS CA745444 434 bp mRNA linear EST 26-NOV-2002
DEFINITION wri2s.pk001.i6 wri2s Triticum aestivum cDNA clone wri2s.pk001.i6 5' end, mRNA sequence.
ACCESSION CA745444
VERSION CA745444.1 GI:25561267
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 434)
AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z., Miao,G., Caraher,N. and Hanafey,M.K.
TITLE DuPont Wheat cDNA Sequence
JOURNAL Unpublished (2002)
COMMENT Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: T7.
FEATURES Location/Qualifiers
source 1. .434
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wri2s.pk001.i6"
/tissue_type="leaf"
/lab_host="DH10B"
/clone_lib="wri2s"
/note="Vector: PGEM-T Easy; Site_1: SmaI; Riband (susceptible) wheat leaves infected with Septoria tritici strain A, 48 hours after infection, subtracted w/ comparable uninfected leaves"

ORIGIN

Query Match 4.7%; Score 100.8; DB 14; Length 434;
Best Local Similarity 67.1%; Pred. No. 8.9e-10;
Matches 141; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

```
Qy       1382 AACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAG 1441
             ||  | | | | | || ||||||| ||| ||||| | | ||  |  ||  | |   | |
Db        381 AAGTGGAAGAGGAAGAGGAAGAGGAGGAAGAAGAAGAAGAAGNNGAGGAAGAAGAGGAGG 322

Qy       1442 AAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAG 1501
             ||    ||||| |||   |  || ||||||      | ||| ||||  | ||||||  ||||
Db        321 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 262

Qy       1502 ACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAG 1561
             | ||||| ||||| |||||||||    |||||||||  ||||||||||  ||||| ||||
Db        261 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 202

Qy       1562 AACCCGAAGAAGAATTGCCGGCAGAAGAAG 1591
             ||   |||||||||      | ||||||||
Db        201 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 172
```

RESULT 4
CA716220/c
LOCUS CA716220 289 bp mRNA linear EST 26-NOV-2002
DEFINITION wdk3c.pk024.i22 wdk3c Triticum aestivum cDNA clone wdk3c.pk024.i22 5' end, mRNA sequence.
ACCESSION CA716220
VERSION CA716220.1 GI:25438013
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 289)
AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z., Miao,G., Caraher,N. and Hanafey,M.K.
TITLE DuPont Wheat cDNA Sequence
JOURNAL Unpublished (2002)
COMMENT Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
FEATURES Location/Qualifiers
source 1. .289
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wdk3c.pk024.i22"
/tissue_type="kernel"
/lab_host="DH10B"
/clone_lib="wdk3c"
/note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2: XhoI; Wheat (Triticum aestivum L.) developing kernel, 14 days after anthesis."
ORIGIN

Query Match 4.7%; Score 99.4; DB 14; Length 289;
Best Local Similarity 67.5%; Pred. No. 1.7e-09;
Matches 139; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

```
Qy       1386 GAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAG 1445
             ||||  |  ||| || ||||  || ||||| |  | ||  |  ||  | |  |||||
Db        252 GAAGATGACGGAGGAGGAGGANGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 193

Qy       1446 CGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGA 1505
             |||||  |||    |  || ||||||      | ||| ||||  | ||||||  ||||| ||
Db        192 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 133

Qy       1506 AGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACC 1565
             |||  |||||  |||||||||   |||||||||  ||||||||||  ||||| |||||
Db        132 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 73
```

```
Qy       1566 CGAAGAAGAATTGCCGGCAGAAGAAG 1591
             |||||||||      | ||||||||
Db          72 AGAAGAAGAAGAAGAAGAAGAAGAAG 47
```

RESULT 5
CA022136
LOCUS CA022136 223 bp mRNA linear EST 23-OCT-2002
DEFINITION HZ42E07r HZ Hordeum vulgare subsp. vulgare cDNA clone HZ42E07 5-PRIME, mRNA sequence.
ACCESSION CA022136
VERSION CA022136.1 GI:24299510
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 223)
AUTHORS Radchuk,V., Zhang,H., Weschke,W., Potokina,E. and Wobus,U.
TITLE Barley ESTs from developing seeds
JOURNAL Unpublished (2002)
COMMENT Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 223 Std Error: 0.00
Plate: 42 row: E column: 7
Seq primer: M13rev.
FEATURES Location/Qualifiers
source 1. .223
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:275541"
/db_xref="taxon:112509"
/clone="HZ42E07"
/tissue_type="pericarp"
/dev_stage="0-7 DAP (days after pollination)"
/lab_host="XL10-Gold"
/clone_lib="HZ"
/note="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of cDNA); Site_2: XhoI (3'-end of cDNA); pericarp 0-7 DAP(days after pollination). Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRIadapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI,PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinats is not 100% reliable.Average insert size is 900 bp"
ORIGIN

Query Match 4.7%; Score 99.2; DB 13; Length 223;
Best Local Similarity 67.3%; Pred. No. 1.8e-09;
Matches 140; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```
Qy       1384 CCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAA 1443
             | | | | | | || ||||| |  || |||| ||  | ||  |  || || |  | |||
Db          1 CGGCACGAGGAAGAAGAAGAAGAAGAAGAAGGGGAAGAAGAAGGGGAAAAAGAAGAGGAA 60

Qy       1444 AGCGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGAC 1503
             |||||  |||    |  || | ||||      |||||  ||||  |  ||||||  |||||
Db         61 GAAGAAGAGGAAGAAGAAGAAGGAGAAGAGGAAGGAGAAGAAGAAGAAGAAGAAGAAGAA 120

Qy       1504 GAAGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAA 1563
             |||||  |||||  |||||||||   |||||||||  ||||||||||  ||||| ||||||
Db        121 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 180
```

```
Qy       1564 CCCGAAGAAGAATTGCCGGCAGAAGAAG 1591
              ||||||||| |   | ||||||||
Db        181 GAAGAAGAAGAAGAGGAAGAAGAAGAAG 208
```

RESULT 6
CE760553
LOCUS CE760553 395 bp DNA linear GSS 30-SEP-2003
DEFINITION tigr-gss-dog-17000369997532 Dog Library Canis familiaris genomic, genomic survey sequence.
ACCESSION CE760553
VERSION CE760553.1 GI:37101172
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 395)
AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
FEATURES Location/Qualifiers
source 1. .395
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site_1: BstXI; Libraries were prepared from peripheral blood"
ORIGIN

Query Match 4.7%; Score 99; DB 29; Length 395;
Best Local Similarity 62.2%; Pred. No. 2e-09;
Matches 156; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
Qy       1357 ATAAAAACCGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAG 1416
              | ||||   |||      |  |   ||    | |   | || ||||| |  || |||||
Db         58 AGAAAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAA 117

Qy       1417 GGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACC 1476
              || | ||| |  ||  | |   |||||   ||||| |||   |  || ||||||
Db         118 GGAGAAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 177

Qy       1477 GGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAA 1536
              | ||| ||||  | |||||| ||||| ||||| ||||| |||||||||   |||||||||
Db         178 GAAGAAGAAGAAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 237

Qy       1537 CCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAGGCAAC 1596
              ||||  |||||  ||||| ||||||   |||||||||       | ||||||||   |
Db         238 GGAGAAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 297

Qy       1597 GGCGGTTCAGG 1607
              |  ||   |||
Db         298 GAAGGAGGAGG 308
```

RESULT 7
BQ241566
LOCUS BQ241566 264 bp mRNA linear EST 03-MAY-2002
DEFINITION TaE05003E07R TaE05 Triticum aestivum cDNA clone TaE05003E07R, mRNA sequence.
ACCESSION BQ241566
VERSION BQ241566.1 GI:20437442
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 264)
AUTHORS Cloutier,S.
TITLE Wheat functional genomics - Glenlea developing seeds cDNA libraries
JOURNAL Unpublished (2002)
COMMENT Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.gc.ca
was cloned directionally, not all sequences generated with reverse primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >2.0 kb
Plate: 003 row: E column: 07
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .264
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TaE05003E07R"
/tissue_type="developing seeds"
/dev_stage="5 days after anthesis"
/lab_host="E. coli DH10B"
/clone_lib="TaE05"
/note="Vector: pSPORT-P (Invitrogen Technologies); Site_1: NotI; Site_2: MluI; mRNA obtained from wheat seeds of cultivar Glenlea 5 days post-anthesis"
ORIGIN

Query Match 4.7%; Score 98.8; DB 13; Length 264;
Best Local Similarity 68.7%; Pred. No. 2.2e-09;
Matches 136; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
              | || ||||| |  || |||||||  | ||  |  ||  | |  ||||||   |||||
Db          43 AAGAAGAAGAAGAAGAAGAAGAGGAAGAAGAAGAAGAAGAAGAAAAAGAAGAAGAAGAAG 102

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
              ||    |  || ||||||      | ||| ||||  | |||||| ||||| ||||| ||||
Db         103 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 162

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
              | |||||||||   |||||||||   ||||||||||  ||||| ||||||   |||||||
Db         163 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 222

Qy       1574 AATTGCCGGCAGAAGAAG 1591
              ||    | ||||||||
Db         223 AAGAAGAAGAAGAAGAAG 240
```

RESULT 8
CE180020
LOCUS CE180020 729 bp DNA linear GSS 25-SEP-2003
DEFINITION tigr-gss-dog-17000326768259 Dog Library Canis familiaris genomic, genomic survey sequence.
ACCESSION CE180020
VERSION CE180020.1 GI:35328914
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

```
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match             4.6%;  Score 98.6;  DB 29;  Length 729;
  Best Local Similarity   67.0%;  Pred. No. 2.5e-09;
  Matches  140;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy       1386 GAAGGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAG 1445
             |||| | || | | || | || |||||  | | || | || | | |||||
Db        168 GAAGAAGAAGAAGAAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 227

Qy       1446 CGAAGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGA 1505
             ||||| |||   |  || ||||||   |  | |||  ||||  | |||||| ||||| ||
Db        228 AGAAGAAGAAGAAGAAGAAGAAGAAGAAAAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 287

Qy       1506 AGATGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACC 1565
             ||| ||||| |||||||||   |||||||||   ||||||||||  ||||| ||||||
Db        288 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGG 347

Qy       1566 CGAAGAAGAATTGCCGGCAGAAGAAGGCA 1594
             ||||||||||||      || || || |||
Db        348 AGAAGAAGAATTAGGATCAAAATAATGCA 376


RESULT 9
CK206118/c
LOCUS       CK206118                1096 bp    mRNA    linear   EST 08-DEC-2003
DEFINITION  FGAS017695 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
            aestivum cDNA, mRNA sequence.
ACCESSION   CK206118
VERSION     CK206118.1  GI:39568508
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 1096)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE     Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [29,696].
            Plate: L5B006  row: F  column: 12.
FEATURES             Location/Qualifiers
     source          1. .1096
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
                     /note="Vector: pCMV.SPORT6; Crown and developmental stages
                     of spike formation in wheat cultivar Norstar. 4 mRNA
                     populations were combined before constructing the library.
                     The first mRNA population is from 1cm crown sections after
                     30 days of cold acclimation. The second is from 1cm crown
                     sections after 11 days of deacclimation (before
                     deacclimation plants were fully vernalized for 49 days).
                     The third is from different developmental stages of spike
                     formation (5 to 50mm) that still have not emerged from the
                     leaf (dissection required). The last is from different
                     developmental stages of spike and seed formation after
                     having emerged from the leaf (visible). First strand
                     synthesis in this library was done in the presence of
                     methylated dCTP thereby protecting from internal cleavage
                     with NotI."
ORIGIN

  Query Match             4.6%;  Score 98.6;  DB 14;  Length 1096;
  Best Local Similarity   63.9%;  Pred. No. 2.6e-09;
  Matches  149;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy       1359 AAAAACCGAACGCCCCGCCGTCCAACCGAAGGCGCAGGATGAAGAGGGGGACGAAGAGGG 1418
             | ||   |||   ||||| |    | | |   |  || ||||| |  || ||||| |
Db        300 AGAACAAGAAGCACCCGCAGCAGCAGCAACAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 241

Qy       1419 TGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACGAAATCGGCGATGAAGAAAGCACCGG 1478
             | ||  |  ||  | |  |||||   ||||| |||   |  || ||||||      |
Db        240 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 181

Qy       1479 AGACGAAGTCGTAGAAGATGAAGACGAAGATGAAGACGAAGAAGAAATCGAAGAAGAACC 1538
             ||| ||||  | ||||||  |||||  |||||  |||||  |||||||||      |||||||||
Db        180 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 121

Qy       1539 TGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAGAATTGCCGGCAGAAGAAG 1591
             | ||||||||  |||||  ||||||   |||||||      || | ||||||||
Db        120 AGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGCCCAACCAGAAGAAGAAG 68


RESULT 10
BZ201507
LOCUS       BZ201507                799 bp    DNA     linear   GSS 11-OCT-2002
DEFINITION  CH230-445O16.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
            CH230-445O16, genomic survey sequence.
ACCESSION   BZ201507
VERSION     BZ201507.1  GI:23859559
KEYWORDS    GSS.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 MboI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-445O16.TV
```

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 445 row: O column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES Location/Qualifiers
source 1. .799
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-445O16"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 4.6%; Score 97.6; DB 28; Length 799;
Best Local Similarity 62.9%; Pred. No. 4e-09;
Matches 151; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```
Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
             | || ||||| |  || |||||  |  | ||  |  ||  | |  |||||  ||||| |
Db         79 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 138

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
             ||   |  || |||||||      | ||| ||||  | |||||| | ||| ||||| ||||
Db         139 AAGAAGAAGAAGAAGAAGAAGATGAAGAAGAAGAGGAAGAAGAAGGAGAAGAAGAAGAAG 198

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
             | |||||||||   |||||||||  ||||||||||  |||||  ||||||   |||||||
Db         199 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 258

Qy       1574 AATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAGCATCCTGCCCACTCCGGAAGCCT 1633
             ||      | ||||||||   | |  ||   ||      |  ||||  ||   |  ||||
Db         259 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGGGGAAGAAGAAAAGTTGCCTACAAGGCTGGCCT 318
```

RESULT 11
BF707975
LOCUS BF707975 575 bp mRNA linear EST 12-JUN-2002
DEFINITION A634_LE Aplysia metacerebral cell cDNA library Aplysia californica cDNA 5', mRNA sequence.
ACCESSION BF707975
VERSION BF707975.1 GI:18002301
KEYWORDS EST.
SOURCE Aplysia californica (California sea hare)
ORGANISM Aplysia californica
Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia;
Anaspidea; Aplysioidea; Aplysiidae; Aplysia.
REFERENCE 1 (bases 1 to 575)
AUTHORS Sadreyev,R.I., Meleshkevich,E.A., Matz,M.V., Panchin,Y.V. and Moroz,L.L.
TITLE Characterization of the gene pool expressed in the individual identified neuron
JOURNAL Unpublished (2001)
COMMENT Contact: Sadreyev RI
The Whitney Laboratory
University of Florida
9505 Ocean Shore Blvd., St Augustine, FL 32080, USA
Tel: 904 461 4044
Fax: 904 461 4008
Email: sadreyev@whitney.ufl.edu
Seq primer: 5' CGACGTGGACTATCCATGAACGCA 3'.

FEATURES Location/Qualifiers
source 1. .575
/organism="Aplysia californica"
/mol_type="mRNA"
/db_xref="taxon:6500"
/tissue_type="nervous system"
/cell_type="Metacerebral cell"
/clone_lib="Aplysia metacerebral cell cDNA library"
/note="Vector: pGEM-T"

ORIGIN

Query Match 4.6%; Score 97.4; DB 10; Length 575;
Best Local Similarity 67.5%; Pred. No. 4.3e-09;
Matches 137; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

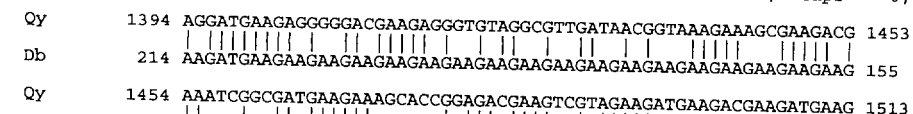
```
Qy       1389 GGCGCAGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGA 1448
             ||  || || ||||| |  || ||||| |  | ||  |  ||  | |  |||||  ||
Db         100 GGGCCAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 159

Qy       1449 AGACGAAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGA 1508
             ||| |||   |  || ||||||      | ||| ||||  | |||||| ||||| |||||
Db         160 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 219

Qy       1509 TGAAGACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGA 1568
              ||||| |||||||||   |||||||||  ||||||||||  ||||| ||||||  ||
Db         220 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 279

Qy       1569 AGAAGAATTGCCGGCAGAAGAAG 1591
             |||||||     | ||||||||
Db         280 AGAAGAAGAAGAAGAAGAAGAAG 302
```

RESULT 12
CG321416/c
LOCUS CG321416 232 bp DNA linear GSS 26-AUG-2003
DEFINITION OGYBI53TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0640J09, genomic survey sequence.
ACCESSION CG321416
VERSION CG321416.1 GI:34238682
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 232)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGYBI53TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES Location/Qualifiers
source 1. .232
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0640J09"

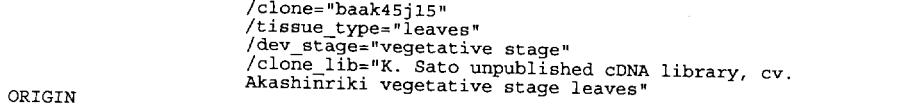
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

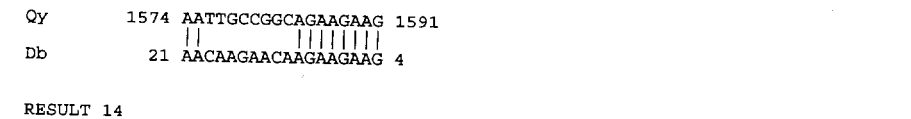
Query Match 4.6%; Score 97.2; DB 29; Length 232;
Best Local Similarity 68.2%; Pred. No. 4.5e-09;
Matches 135; Conservative 0; Mismatches 63; Indels 0; Gaps 0;



RESULT 13
BJ454017/c
LOCUS BJ454017 233 bp mRNA linear EST 23-MAY-2002
DEFINITION BJ454017 K. Sato unpublished cDNA library, cv. Akashinriki vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone baak45j15 5', mRNA sequence.
ACCESSION BJ454017
VERSION BJ454017.1 GI:21132592
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 233)
AUTHORS Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .233
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baak45j15"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv. Akashinriki vegetative stage leaves"
ORIGIN

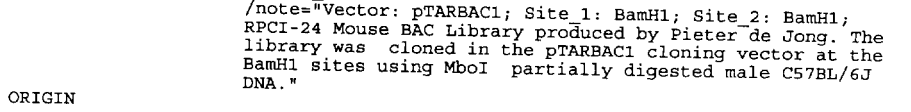
Query Match 4.6%; Score 97.2; DB 12; Length 233;
Best Local Similarity 68.2%; Pred. No. 4.5e-09;
Matches 135; Conservative 0; Mismatches 63; Indels 0; Gaps 0;





RESULT 14
BH040136/c
LOCUS BH040136 691 bp DNA linear GSS 17-JUL-2001
DEFINITION RPCI-24-255A23.TV RPCI-24 Mus musculus genomic clone RPCI-24-255A23, genomic survey sequence.
ACCESSION BH040136
VERSION BH040136.1 GI:14818806
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 691)
AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M., Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E., Russell,D., de Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-24
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: RPCI-24-255A23.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 255 row: A column: 23
Seq primer: T7
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .691
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-255A23"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA."
ORIGIN

Query Match 4.6%; Score 97.2; DB 28; Length 691;
Best Local Similarity 68.2%; Pred. No. 4.8e-09;
Matches 135; Conservative 0; Mismatches 63; Indels 0; Gaps 0;



```
                ||   |  || |||||| |    ||||  ||||  | |||||| ||||| ||||| ||||
Db        415 AAGAAGAAGAAGAAGAATGAGGAGGAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 356

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
                | |||||||||   |||||||||   ||||||||||  |||||  |||||||  |||  |||
Db        355 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAACAAGAACAAG 296

Qy       1574 AATTGCCGGCAGAAGAAG 1591
                ||       | ||||||||
Db        295 AAGAAGAAGAAGAAGAAG 278


RESULT 15
BZ104011
LOCUS       BZ104011                675 bp    DNA     linear   GSS 10-OCT-2002
DEFINITION  CH230-227F5.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-227F5, genomic survey sequence.
ACCESSION   BZ104011
VERSION     BZ104011.1  GI:23744890
KEYWORDS    GSS.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 EcoRI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: CH230-227F5.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/or ering_information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
            Plate: 227  row: F  column: 5
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-227F5"
                     /sex="Female"
                     /cell_type="Brain"
                     /clone_lib="CHORI-230 Segment 1"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
ORIGIN

  Query Match             4.6%;  Score 97;  DB 28;  Length 675;
  Best Local Similarity   65.4%;  Pred. No. 5.2e-09;
  Matches  142;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy       1394 AGGATGAAGAGGGGGACGAAGAGGGTGTAGGCGTTGATAACGGTAAAGAAAGCGAAGACG 1453
                | || |||||||  || |||||||  |  |  |  ||  | |   | |||   ||||| |
Db        319 AAGAAGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGGAAGAGG 378

Qy       1454 AAATCGGCGATGAAGAAAGCACCGGAGACGAAGTCGTAGAAGATGAAGACGAAGATGAAG 1513
                ||   |  || ||||||      | ||| ||||  | |||||| ||||| ||||| ||||
```

```
Db        379 AAGAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 438

Qy       1514 ACGAAGAAGAAATCGAAGAAGAACCTGAAGAAGAAGCTGAAGAGGAAGAACCCGAAGAAG 1573
                | |||||||   |||||||||   ||||||||||  |||||  |||||||  |||||||
Db        439 AAGAAGAAGAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 498

Qy       1574 AATTGCCGGCAGAAGAAGGCAACGGCGGTTCAGGCAG 1610
                ||       | |||||||||  |    | ||   | | ||
Db        499 AAGAAGAAGAAGAAGAAGGAAGAAGAGGAAGAAGAAG 535


Search completed: August 25, 2004, 05:21:14
Job time : 4848.67 secs
```